MATHEMATICS MAGAZINE

Vol. 25, No. 2, Nov.-Dec., 1951

CONTENTS

OUR CONTRIBUTORS

Clarence DeWitt Smith, Associate Professor of Statistics, University of Alabama, was born in Jackson, Mississippi in 1891, and received his college education at Mississippi College (A.B., '15) and at the Universities of Chicago (A.B., '16) and Iowa (M.S., '25, Ph.D., '27). After a few years as a teacher and Principal in the high schools of Mississippi and Louisiana, Dr. Smith served as an artillery officer during World War I. After the war, he was appointed Professor of Mathematics and Department Head at Louisiana College. Later he held similar posts at Mississippi A. & M. College and at Mississippi State College. He was appointed to his present position in 1948. Prof. Smith's principal mathematical interests are statistics and geometry. He has been on the editorial board of this Magazine for many years.

John E. Freund, Associate Professor of Mathematics, Alfred University, was born in 1922. After attending the University of London (England) and U.C.L.A. (B.A., '43, M.A., '44), he worked for a year as a research statistician for the Mutual Casualty Insurance Rating Bureau (N.Y.), and then joined the faculty of Alfred University in 1946. His main interests are mathematical statistics, probability, logic and the philosophy of science.

Homer V. Craig, Professor of Applied Mathematics at the University of Texas, was born in Denver, Colorado in 1900. He attended the University of Colorado (A.B., '24) and the University of Wisconsin (Ph.D.,'29), and then joined the faculty of the University of Texas. His research interests are tensor analysis and related subjects. Professor Craig is also known as the author of a textbook on vector analysis. He has been an editor of this Magazine since 1947.

Gene F. Rose, Instructor in Mathematics, University of Wisconsin, was born in Erie, Pennsylvania in 1918. He studied at Case (B.S.,'38), Northwestern, Carnegie Tech. and Pittsburgh in Physical Science and since 1945 has specialized in the foundations of mathematics at Wisconsin (M.S.,'47). He expects to receive the doctorate soon. Mr. Rose's article, "A Proof Schema for a Class of Theorems", appeared in the September-October issue.

S. E. Urner, Professor of Mathematics and Department Chairman, Los Angeles State College of Applied Arts and Sciences, studied at Baker University (A.B.,'06) and Harvard University (Ph.D., '11). In 1918 Dr. Urner left the teaching profession and became a fruit farmer in the Santa Clara Valley and later a business man in Bakersfield. In 1931 he returned to teaching and was appointed to the Faculty of Los Angeles City College, where he became Department Chairman in 1946. He has held his present position since 1950. Prof. Urner (with the late William B. Orange) is the co-author of two textbooks, the second of which (Elements of Mathematical Analysis, 1950) embodies the ideas expressed in his article, "Unified Freshman Mathematics" (September-October, 1951).

SOME PROBABILITY ESTIMATES FROM CONTINGENCY TABLES

C. D. Smith

The problem of comparing joint outcomes from two samples with equal number N arises when it is desirable to arrange the data in mutually exclusive classes. In such cases the order of joint outcomes may be arranged in a contingency table. Karl Pearson calculated a coefficient of contingency and also applied the chi-square test. Wilks and others have given probability distributions for r-successes in samples of N, $r < N$, where a success is a mated pair of outcomes. For example, two samples of corresponding items are each classes as good (G), or bad (B). A success is indicated by the joint outcomes GG or BB. The order of all joint outcomes may be recorded in the contingency table

	G	B
G	GG	GB
B	BG	BB

Assuming that G has probability p and B has probability q the binomial $(p + q)^2 = p^2 + 2pq + q^2$ gives the order of the cell numbers where $2pq$ is composed of pq and qp. If $p = q$ the cell numbers in the table are equal and the table may be represented with the number x in each cell. For $p \neq q$ the cell numbers are not equal and it will be convenient to use the general notation as follows:

	G	B
G	x_{ij}	x_{ij}
B	x_{ij}	x_{ij}

The subscripts i and j indicate the respective rows and columes. Represent the sums of rows and columes by R_i and C_j respectively. The number of successes is given by $L = x_{11} + x_{22}$. The weighted average number of all joint outcomes is given by $A = (C_1R_1 + C_2R_2)/N$, where the number of a given outcome from the first sample is weighted with the number of the corresponding outcome from the second sample.

We define a measure of relative discrepancy by

$$f = (L - A)/N. \tag{1}$$

The characteristics of the measure of discrepancy are given by the

following results. For $x_{11}/x_{21} = x_{12}/x_{22}$, it follows that $x_{11} = C_1R_1/N$, $x_{22} = C_2R_2/N$, and $L = A$ which gives $f = 0$. When the cell numbers are not proportional, $L \neq A$ and f is either positive or negative. In this case for $x_{11}/x_{21} \gtrless x_{12}/x_{22}$, $L \gtrless A$ and $f \gtrless 0$. The results mean that f is a measure of the failure of the table to give proportional numbers of joint outcomes. One extreme case occurs when all joint outcomes give the equal numbers $GG = BB$. In this case $f = ½$. In the other extreme all outcomes give equal numbers $BG = GB$. In this case $f = -½$. When f is calculated for all possible outcomes in samples of eight the distribution of f with interval .1 is approximately normal with mean zero and $\sigma = .2$. Under these conditions we may assume that the distribution of f is approximately normal with mean zero and $\sigma = .2$, for N not less than eight.

In applying the method, the joint outcomes in two samples are arranged in a contingency table, the value of f is calculated, and the corresponding probability taken from the normal table of areas. A problem taken from the field of industrial management illustrates the practical value of the method. On the basis of aptitudes tests a group of 100 applicants for positions as filing clerks are rated as good or bad. After due trial on the job each individual is classed by the record as good or bad. The question for management is the extent to which the tests may indicate expected success on the job. For a given case we have the table

	G	B	
G	45	20	65
B	10	25	35
	55	45	

$L = 70$, $A = 51.5$, and $f = (L - A)/N = .185$. From the table of areas at $x/\sigma = .92$ the probability is $P = .3212$. Since the probability for improving the test estimates is given by increasing f one would expect better results to be possible in about 18% of all cases. In some cases this result may be considered good enough while in others it may be desirable to seek a better method of rating applicants.

Other applications arise in quality inspection and control. The output from each of a group of workers or from machines may be classified both before and after a program for improvement has been carried out. Two inspections may be compared to determine the extent to which machines are losing efficiency through wear. In physics a method of prediction may be compared with later observed results to determine the value of prediction in such cases as, breaking strength, crushing weight, occurrence of daily rainfall, or daily maximum temperatures. These problems are mentioned to indicate the wide variety of important problems which may be studied by use of the contingency table and the value of

the relative discrepancy f, where an estimate of the probability of better correspondence is desired.

The case for three mutually exclusive events has some interesting variations from that of two events. The classes may be represented by good (G), acceptable (A), and bad (B). With probabilities p_1, p_2, p_3, of respective outcomes G, A, B, in each sample, the order of joint outcomes is given by the expansion of $(p_1 + p_2 + p_3)^2 = p_1p_1 + p_2p_2 + p_3p_3 + p_1p_2 + p_2p_1 + p_1p_3 + p_3p_1 + p_2p_3 + p_3p_2$. The table of joint outcomes is as follows:

	G	A	B
G	GG	GA	GB
A	AG	AA	AB
B	BG	BA	BB

For cell numbers of joint outcomes x_{ij}, we have for sums of rows R_i, for sums of columes C_j, and the sample number is N. A model experiment for the case of equal probabilities may be arranged as follows. Use two triangular regular prisms with one letter on each face. Place pegs in each end so that each, when thrown, will come to rest on one face. Throw the two prisms and take the letters on which they lie as a joint event. Repeat the experiment N times and record results in the contengency table.

In the table of cell numbers x_{ij} it turns out that $x_{11} = C_1R_1/N$, and similar results for x_{22} and x_{33} provided that $x_{11}/x_{21} = x_{12}/x_{22} = x_{13}/x_{23} = x_{21}/x_{31} = x_{22}/x_{32} = x_{23}/x_{33}$. The number of like outcomes, or mated pairs, is $L = x_{11} + x_{22} + x_{33}$ and the weighted average of all joint outcomes is $A = \Sigma C_jR_i/N$. Here as before the sum of each row is weighted with the sum of the corresponding colume. For proportional cell numbers as given above it is true that $L = A$, otherwise $L \neq A$. One extreme occurs when all joint outcomes are mated pairs such that $x_{11} = x_{22} = x_{33}$. Here $f = (L - A)/N = 2/3$. The other extreme occurs when all joint outcomes fall in the cells $x_{13} = x_{31}$. In this case $f = -1/2$. The case of proportional cell numbers has been stated for which $f = 0$. It is now evident that for three mutually exclusive categories, the values of f are distributed about $f = 0$ with range -1/2 to 2/3. Obviously the distribution of values of f is not normal and estimates of the probability must be taken from a suitable frequency curve. In practical applications the use of Pearson's Type III frequency curve has given satisfactory estimates for the probability of a given value of f.

To illustrate a case of three mutually exclusive events, the following table represents 200 placement test scores paired with the record of the same 200 freshmen in algebra.

		Test Score G	A	B	
Algebra Score	G	25	4	0	29
	A	5	80	13	98
	B	0	11	62	73
		30	95	75	200

From the table, $f = (L - A)/N = .44$. In Pearson's Type III with the origin at the mode, the probability of a deviation greater than .44 on the range extending to .67 is fairly small, and the placement test may be considered a good estimate of expected performance in algebra.

In general the table for n-classes gives the positive extreme value of $f = (n-1)/n$ when all pairs are mated such that $x_{11} = x_{22} = \ldots = x_{nn}$, in which $L = N$. The extreme negative value $f = -1/2$ occurs when all pairs are mismated such that $x_{1n} = x_{n1}$, in which case $L = 0$. Evidently the degree of positive skewness increases as the number of classes increases. For $n > 3$, a suitable curve from the Gram-Charlier system should give good estimates of the probability that values of f may be greater than a given value.

University of Alabama

COLLEGIATE ARTICLES

Graduate training not required for reading

THE TRANSFER DISTRIBUTION

John Freund

An important type of probability sequence is the sequence where the probability of each element (trial) is determined exclusively by the result of the immediately preceding trial. These sequences, which belong to the class which is commonly referred to as simple stationary Markoff chains, are called *sequences with probability transfer*.[1] It is our purpose to develop in this paper the transfer distribution, i.e. the probability of obtaining "x successes in n trials," if the probability of a success depends exclusively on the outcome of the immediately preceding trial.

Let us denote a "success" by the attribute B and a "failure" by the attribute $\overline{B}$ (non-B). We shall consider as known, the following four probabilities which are defined as:

i) the probability that a B will be succeeded by a B, or $P(B,B') = p_{11}$.

ii) the probability that a B will be succeeded by a $\overline{B}$, or $P(B,\overline{B}') = p_{12}$.

iii) the probability that a $\overline{B}$ will be succeeded by a B, or $P(B,B') = p_{21}$.

iv) the probability that a $\overline{B}$ will be succeeded by a $\overline{B}$, or $P(\overline{B},\overline{B}') = p_{22}$.

It can easily be shown that the probability $P(B)$, i.e., the probability of obtaining a B, regardless of the predecessor, is

$$P(B) = \frac{p_{21}}{1 - p_{11} + p_{21}} \tag{1}$$

which follows directly from the equation:

$$P(B) = P(B)P(B,B') + [1 - P(B)]P(\overline{B},B') \tag{2}$$

In order to simplify our notation, we shall write $P(B)$ simply as p.

The number of ways (combinations) in which we can arrange x B's and $(n - x)$ $\overline{B}$'s is $\binom{n}{x}$. In contrast to the Binomial distribution, where the probability of obtaining any one of these arrnngements is the same, we find that for transfer sequences, the probabilities which are assoc-

[1]This terminology is due to H. Reichenbach, "The theory of probarility," University of California Press, 1949.

iated with the various combinations depend on two factors. First of all, the probability will depend on whether the first element is a B or a $\overline{B}$ and secondly it will depend on the number of runs[2] which appear in the arrangement. In order to find the probability of obtaining x B's in n trials, we shall use the following well known theorems from the field of combinatorial statistics:[3]

A) If we have x elements of one kind and $(n - x)$ of another, then the number of ways in which they can be arranged to form $2k$ runs is

$$\binom{x-1}{k-1}\binom{n-x-1}{k-1} \tag{3}$$

B) The number of ways in which they can be arranged to form $2k + 1$ runs if the first element is a B, is

$$\binom{x-1}{k}\binom{n-x-1}{k-1} \tag{4}$$

and if the first element is a $\overline{B}$, it is

$$\binom{x-1}{k-1}\binom{n-x-1}{k} \tag{5}$$

In the above expressions the variable k goes from 1 to x if $x \leq n - x$. In order to simplify our expressions, we shall therefore assume that $x \leq n - x$, which does not offer any undue restrictions, because we can always calculate the probability of obtaining $(n - x)$ $\overline{B}$'s if $n - x < x$ by interchanging the p_{ij} in our final results.

If we have an even number of runs (namely $2k$) and if the first element is a B, the following must be true in regard to the combination of the attributes:

i) there are k $\overline{B}$'s which are preceded by B's
ii) there are $k - 1$ B's which are preceded by $\overline{B}$'s
iii) there are $n - x - k$ $\overline{B}$'s which are preceded by $\overline{B}$'s
iv) there are $x - k$ B's which are preceded by B's.

The probability of obtaining x B's and $n - x$ $\overline{B}$'s which form $2k$ runs and where the first element is a B, is therefore

$$p_1 = \binom{x-1}{k-1}\binom{n-x-1}{k-1} p_{12}^{k} p_{21}^{k-1} p_{22}^{n-x-k} p_{11}^{x-k} p$$

[2]A run is a sequence of letters of the same kind which is bounded by letters of the other kind, or which is at the beginning or the end of the total sequence.

[3]See S. S. Wilks, "Mathematical Statistics," Princeton Univ. Press 1944, p. 203.

in which any binomial coefficient $\binom{a}{b} = 0$, if $a < b$. If the first element is a $\bar{B}$, the corresponding probability is

$$p_2 = \binom{x-1}{k-1}\binom{n-x-1}{k-1} p_{12}^{k-1} p_{21}^{k} p_{22}^{n-x-k} p_{11}^{x-k}(1-p) \qquad (7)$$

If we have an odd number of runs $(2k+1)$ and the first element is a B, the following must be true for the sequence of attributes:

i) there are k $\bar{B}$'s which are preceded by B's
ii) there are k B's which are preceded by $\bar{B}$'s
iii) there are $n - x - k$ $\bar{B}$'s which are preceded by $\bar{B}$'s
iv) there are $x - k - 1$ B's which are preceded by B's

The probability of obtaining x B's and $(n - x)$ $\bar{B}$'s, which form $2k + 1$ runs and where the first element is a B, is therefore

$$p_3 = \binom{x-1}{k}\binom{n-x-1}{k-1} p_{12}^{k} p_{21}^{k} p_{22}^{n-x-k} p_{11}^{x-k-1} p \qquad (8)$$

The corresponding probability, if the first term is a $\bar{B}$, is

$$p_4 = \binom{x-1}{k-1}\binom{n-x-1}{k} p_{12}^{k} p_{21}^{k} p_{22}^{n-x-k-1} p_{11}^{x-k}(1-p) \qquad (9)$$

The total probability of obtaining x B's in n trials (if $x \leq n - x$ and if x does not equal 0) is the sum of these four probabilities, summed on k, i.e.

$$f(x) = \sum_{k=1}^{x} (p_1 + p_2 + p_3 + p_4) \qquad (10)$$

or

$$f(x) = \sum_{k=1}^{x}\left[\binom{x-1}{k-1}\binom{n-x-1}{k-1} p_{12}^{k} p_{21}^{k-1} p_{22}^{n-x-k} p_{11}^{x-k} p + \binom{x-1}{k-1}\binom{n-x-1}{k-1} p_{12}^{k-1} p_{21}^{k} p_{22}^{n-x-k} p_{11}^{x-k}(1-p) + \binom{x-1}{k}\binom{n-x-1}{k-1} p_{12}^{k} p_{21}^{k} p_{22}^{n-x-k} p_{11}^{x-k-1} p + \binom{x-1}{k-1}\binom{n-x-1}{k} p_{12}^{k} p_{21}^{k} p_{22}^{n-x-k-1} p_{11}^{x-k}(1-p)\right]. \qquad (11)$$

Substituting (1) into expression (11), we finally obtain

$$f(x) = \frac{1}{1 - p_{11} + p_{21}} \sum_{k=1} \left\{ 2\binom{x-1}{k-1}\binom{n-x-1}{k-1} p_{12}^{k} p_{21}^{k} p_{22}^{n-x-k} p_{11}^{x-k} \right.$$
$$+ \binom{x-1}{k}\binom{n-x-1}{k-1} p_{12}^{k} p_{21}^{k+1} p_{22}^{n-x-k} p_{11}^{x-k-1} \tag{12}$$
$$\left. + \binom{x-1}{k-1}\binom{n-x-1}{k} p_{12}^{k+1} p_{21}^{k} p_{22}^{n-x-k-1} p_{11}^{x-k} \right\}$$

If x equals zero, we can obtain directly

$$f(0) = \frac{p_{12}p_{22}^{n-1}}{1 - p_{11} + p_{21}} \tag{13}$$

It can easily be shown that in the special case where $p_{11} = p_{21}$ expression (12) will reduce to the well known formula for the binomial distribution.

If n is large the Transfer Distribution approaches a normal distribution which has the mean

$$m = \frac{np_{21}}{1 - p_{11} + p_{21}} \tag{14}$$

and whose standard deviation is approximately equal to

$$\sigma = \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{1 \pm \frac{2(p_{11} - p)}{1 - p_{11}}} \tag{15}$$

where p stands for the probability $P(B)$ which is given by expression (1). Formula (15) is not an exact formula, but an approximation which was derived by V. Bargmann[4]. If $p_{11} = p_{21} = p$ then (14) and (15) reduce to the corresponding moments of the ordinary binomial distribution.

[4]See H. Reichenbach, "The Theory of Probability," p. 290.

Alfred University

VECTOR ANALYSIS

Homer V. Craig

1. **The nature and history of vector analysis.** For the purpose of orientation, vector analysis may be regarded quite advantageously (however somewhat inaccurately) as a generalization of the algebra and calculus of real numbers resulting from the replacement of the real number as the basic entity by the concept directed stroke or vector regarded as embodying the couplet of ideas *magnitude* and *direction*. Such a brief characterization is not completely satisfactory because it is an oversimplification from both the historical and the conceptual viewpoints. Historically, the situation is somewhat complex. Vector analysis as a mathematics of directed strokes was extracted from two more extensive theories – the Ausdehnungslehre of Grassmann (1840-44) and Hamilton's Quaternion Analysis (1843-1844) – two works of first rank historical importance. Conceptually, the term *vector* is applied to more abstract entities than directed line segments in ordinary three-dimensional space. The spaces involved may be N-dimensional with N an unspecified integer, or they may have infinitely many dimensions. The basic patterns for these higher vector theories, however, derive quite naturally from the same parent theory, the elementary vector analysis based on the directed stroke. Thus we have on the one hand an elementary theory extracted from more comprehensive and profound theories but rooted in synthetic three-dimensional geometry, while on the other we have a contrasting theory (or set of theories) based on natural generalizations of the dominant algebraic structures of the elementary theory. To sum up, the vector analysis of directed strokes was an offshoot of the *theory of quaternions* and the remarkably inclusive, multi-dimensional *Ausdehnungslehre* – two revolutionary works which mark the beginning of a new epoch in algebra (both presented among other things the startling innovation of products of quantities such that $ab \neq ba$).

Mathematics grows by virtue of certain processes which in degree at least are peculiar to it and the exact sciences. Among these, and of special interest to us at the moment because it is well exemplified by vector analysis, is the process of *creative extension* or *modification* of key entities, sometimes called *generalization* because the resulting theory may embrace the original as a special case. Such generalizations although they may seem simple in retrospect are evidently not easy to make, but once accomplished growth may be accelerated to an astounding pace. Some creative modifications take the form of questioning the *urgency*, *efficacy*, or *validity* of sanctioned viewpoints and assumptions. For example, one of the major developments of vector algebra stems from a denial of the urgency of the *commutative law of multiplication*. E. T. Bell, commenting on Hamilton's quaternion theory and its non-

commutative multiplication, states: "It is radical departures from traditional orthodoxy such as these that carry mathematics forward what seems like a century or more at one stride."

As an instance of creative generalization, let us consider the step from *arithmetic* to *algebra*. In elementary arithmetic, we are given a set of abstract objects, *the numbers*, and two basic processes addition and multiplication. These processes are such that from *each* given pair of numbers there is determined a unique number called their *sum* and a second unique number called their *product*. In our work with these processes and their inverses (subtraction, division, and the extraction of roots), we are concerned almost exclusively with *specific* numbers. By simply replacing the specific numbers with letters denoting *unspecified* numbers we open the door to *algebra*, a subject which is at once vastly more extensive and vastly more powerful than the parent *elementary arithmetic* from which it arose. So long as the fundamental entity denoted by the letters is taken to be *number* (a single real number rather than a set of real numbers or something more abstract), this generalization is essentially of a purely logical nature.

2. **Vectors and their sums.** *Vector analysis*, however, furnishes a generalization of an entirely different type. This generalization is obtained by replacing the entity *number* with a more complicated concept - the *directed stroke*, which involves the couplet of ideas *number* and *direction*. The *number* associated with the directed stroke is its *length* in terms of a given unit and the correlated *direction* is, of course, the direction of the stroke. Thus the overall program of elementary vector analysis is to create a *mathematics of directed strokes* in contradistinction to the mathematics of *real numbers*. The motivation for much of the development is supplied by elementary mechanics - a subject abounding in quantities involving direction such as: displacement, force, velocity, acceleration, etc.

The earlier writers on quaternions, and Gibbs (Elements of Vector Analysis 1881, 1884; Velocity of Propagation of Electrostatic Force 1896) have used Greek letters for vectors and reserved Latin letters for scalars, while other writers have done just the reverse. Still other schemes are to denote vectors by capitals or by letters with superposed bars. We shall follow the practice of Gibbs as set forth in the following quotation "As it is convenient that the form of the letter should indicate whether a vector or a scalar is denoted, we shall use the small Greek letters to denote vectors, and the small English letters to denote scalars. (The three letters i, j, k will make an exception...)."

Vector Addition. Our first problem is the formulation of the definition of *the sum of two vectors* or directed strokes. What meaning shall we ascribe to the phrase *the sum of two vectors*? Before attempting to answer this question, let us examine some of the physical counterparts of vectors. The first and perhaps simplest is the idea of *straight line displacement* - a particle or object is displaced in a straight

line from the point A to the point B. We shall correlate this displacement to the symbol $\vec{AB}$, which represents the directed stroke beginning at A and ending at B. We regard the essential part of the displacement as the fact that the particle was at A and now is at B. If we follow this given displacement with a second displacement say the displacement from B to C, then the single displacement which is equivalent to the two separate displacements $\vec{AB}$ and $\vec{BC}$, in the sense that it takes the particle from A to C, is, of course, the displacement $\vec{AC}$. This suggests

DEFINITION (2.1) $\vec{AB} + \vec{BC} = \vec{AC}$

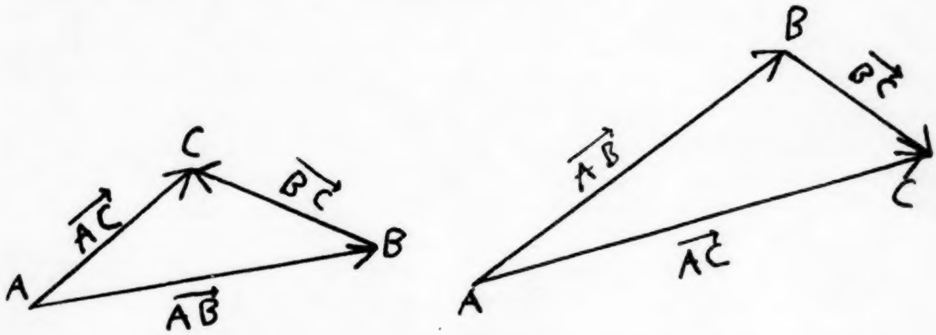


Fig. (2.1)

With regard to this definition, the reader should bear in mind constantly two facts: (a) Definitions are logically arbitrary and are not subject to proof or disproof; (b) We are not defining the sum of the lengths of two strokes. The length of a stroke is a real *number*, and thus the sum of the lengths of two strokes is merely the sum of two real numbers and belongs to the realm of the mathematics of the real number system. We observe that since the straight line furnishes the shortest path between two points, the sum of the mere lengths of the vectors $\vec{AB}$ and $\vec{BC}$ is greater than or equal to the length of the vector sum $\vec{AC}$. The equality holds only in case the three points A, B, C, lie on the same line.

A second instance of a vector quantity is furnished by the physical concept *force*. If a force is applied to a particle the force cannot be described completely by specifying its *magnitude* alone. It develops however, that in this case the force is adequately described by giving in addition to its magnitude, its direction of application, i.e., by specifying its *vector*. Furthermore, it is found that the single force which is equivalent to two given forces α and β acting on a particle is the force corresponding to the vector sum of α and β as given in the preceding definition. This means that if α and β are two forces acting simultaneously on a particle, then $-(\alpha + \beta)$ is the single force which will wipe out the combined effect of the other two. In computing $\alpha + \beta$, we move β by parallel displacement until its origin is at the vertex

or terminus of α and then draw the vector $\alpha + \beta$ from the origin of α to the new terminus of β.

We now *define* two unrestricted vectors (i.e. unrestricted as to locality) to be *equal* if they have the same magnitude and the same direction. Thus, unless we have previously agreed to keep a given vector *localized*, we are free to move it as we please so long as we do not change its *magnitude* or *direction*. The equality of vectors (the symbol is =) as here defined is evidently less stringent than identity. The fact that our definition of the sum of two vectors is applicable to both displacements and forces augurs well for a wide applicability.

We now turn to a second quantity that can be described by giving its *magnitude* and *direction*; namely, *velocity*. If a particle is moving from south to north, for example, with a speed of say ten centimeters per second, this information can be specified by means of a vector ten units in length extending from south to north. Thus *velocities* may be adequately described by means of *directed strokes* or *vectors*. Furthermore, the assumption that velocities add vectorially is apparently in agreement with physical reality. To examine a concrete instance, let us suppose that a ball rolls westwardly along an east-west line drawn on a table with a speed of four centimeters per second relative to the table. Further, let us suppose that while this is happening the table is moving northward three centimeters per second relative to the floor. Relative to the floor, the ball is displaced four centimeters west and three centimeters north of its original position in one second. This displacement is in exact agreement with that determined by adding the two velocity vectors according to the rule given.

The vector addition of velocities has an important application to the navigation of aircraft. The fundamental formula is:

The velocity of the plane relative to the ground is the vector sum of the velocity of the plane relative to the air and the velocity of the air relative to the ground.

Algebraic properties of the sum and difference of two vectors. We have defined the sum of two vectors α and β to be the vector obtained by placing the beginning or origin of β at the terminus of α and drawing the stroke from the origin of α to the new terminus of β. In moving β it is understood that neither the magnitude nor the direction of β is altered. We denote the sum of two vectors α and β by the symbol $\alpha + \beta$. The symbol + used in this way does not have the same meaning as the + in 2 + 3 since the entities involved α, β and 2, 3, are entirely different in the two cases. Furthermore, as we have already pointed out, the length of $\alpha + \beta$ is usually less than $a + b$, the sum of the lengths of α and β. It is imperative that we keep in mind the distinction between $\alpha + \beta$ and $a + b$. Simple geometric arguments will show that (a) $\alpha + \beta = \beta + \alpha$, the *commutative law of addition*, and (b) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$, *the associative law of addition*. This last relationship becomes fairly obvious if we select a point A for the origin of α, let β be the

unique point such that $\overrightarrow{AB} = \alpha$, C and D the unique points such that $\overrightarrow{BC} = \beta$ and $\overrightarrow{CD} = \gamma$. The relationship (*b*), then assumes the form $(\overrightarrow{AB} + \overrightarrow{BC}) + \overrightarrow{CD} = \overrightarrow{AB} + (\overrightarrow{BC} + \overrightarrow{CD})$. The effect of the addition in $\overrightarrow{AB} + \overrightarrow{BC}$, for example, is to leave out the intermediate point B and go directly from A to C. Thus both members of the relationship in question by leaving out the repeated letters reduce to $\overrightarrow{AD}$. Evidently, this law can be extended to the addition of more than three vectors and should be retained in the postulational construction of a vector space of higher dimensions. To recapitulate, in the addition of vectors the *order* and *grouping* may be selected as we please.

The difference of two vectors $\beta - \alpha$ is defined to be the vector γ which satisfies the equation $\alpha + \gamma = \beta$, thus $\beta - \alpha$ is the vector having the property that if it is added to α the result will be β, or in symbols, $\alpha + (\beta - \alpha) = \beta$. This vector $\beta - \alpha$ may be obtained by drawing α and β from a common origin and then drawing from the terminus of α to the terminus of β (see figure (2.2). It may be helpful in a specific case to observe the bearing of the vectors. Thus in the figure β extends

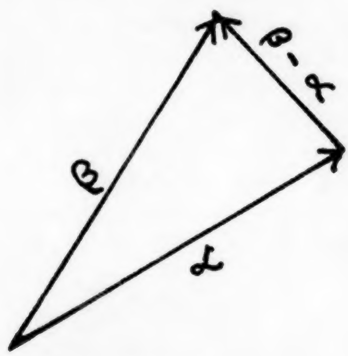


Fig. (2.2)

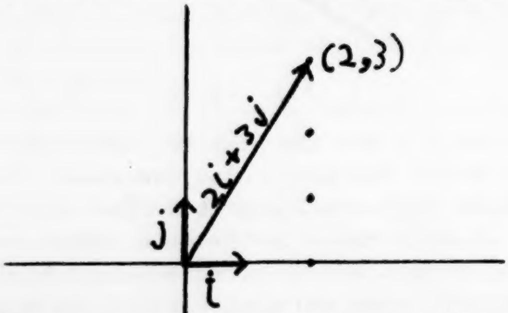
upward more than α while β has the greater rightward component, thus $\beta - \alpha$ bears upward and to the left. Also, we should observe that our figure would collapse if β were equal to α. The symbol $\alpha - \alpha$ is also denoted by 0, and we refer to it as the *zero vector* and regard it as representing a degenerate vector of zero length and indeterminate direction.

3. **Multiplication of a vector by a scalar.** If x is any positive real number and α any vector then the symbol $x\alpha$ is by *definition* taken to represent the vector having the direction of α but having magnitude of α. Thus 2α is a vector having twice the magnitude of α but the same direction. The symbolic products 0α and $(-1)\alpha$ are *defined* to be the *zero vector* and the vector having the magnitude of α but *oppositely directed*, respectively. We also make the definition $-\alpha = (-1)\alpha$. We may now define multiplication of a vector by a negative number $-x$ by

means of the relationship $(-x)\alpha = -(x\alpha)$ thus $-x$ times α is merely the reversal of x times α. To recapitulate, multiplication of a vector by a positive number other than unity stretches or contracts it, while multiplication by a negative number has the added effect of bringing about a reversal.

Incidentally, the definition of multiplication of a vector by a negative number gives an interesting approach to the law of signs of elementary algebra. Thus $(-1)[(-1)\alpha]$ is the reversal of the reversal of α which is simply α itself. Now since we desire to have $(-1)[(-1)\alpha] = [(-1)(-1)]\alpha$, we should define $(-1)(-1)$ to be $+1$, and similarly $(-x)(-y)$ should be xy.

4. **Cartesian coordinates.** A coordinate system in a plane is essentially a scheme for attaching ordered number pairs to points. A point 0 in a plane together with a set of two unit vectors i and j lying in the plane and mutually perpendicular constitute a basis for the familiar rectangular Cartesian coordinate system of elementary analytic geometry. The correlation between the number pairs (a,b) and the points is achieved by the agreement that the number pair (a,b) is to be associated with the *terminal* point P of the vector $ai + bj$ when this vector is drawn from the origin 0. The numbers a and b are called the *coordinates* of P and also the *components* of the vector $ai + bj$. Thus to locate the point corresponding to the number pair $(2,3)$ we start at the origin and generate the vector $2i$ by moving two units in the i-direction, then from this new position we turn through a right angle and move three units in the j-direction. This last displacement adds $3j$ to $2i$ and hence our terminal point is the terminus of the vector $2i + 3j$ when this vector is drawn from the origin. If we agree to let ρ be a *radius vector*, i.e., a vector whose origin is anchored at the



origin of coordinates, o, then we may say that the point corresponding to $(2,3)$ is the endpoint or terminus of the vector ρ given by $\rho = 2i + 3j$. Thus the *first* number of the number pair is the multiplier of the *first* vector of the set i, j. Similarly, the point corresponding to -2, 3 is the terminus of ρ for $\rho = -2i + 3j$. Multiplication of a vector by a negative number reverses it, hence $-2i$ points to the left if i points to the right. Therefore, $(-2,3)$ or more properly the point it represents, lies to the left of the origin.

In addition to the assumption that we have given the plane, the point 0, and the unit vectors $\boldsymbol{i}$ and j we assume that: (a) given any two distinct points in the plane A, B there exists a unique vector AB joining A and B; (b) any vector in the plane admits of a unique representation of the form $ai + bj$, and (c) any point A and any vector α determine a unique point B such that $\vec{AB} = \alpha$. These assumptions are sufficient to ensure that each number pair (a, b) determines a unique point and conversely each point has a unique number pair correlated to it.

It is customary to call the line through 0 containing the vector i the *x-axis*, and the line through 0 containing j, the *y-axis*. The coordinates of an unspecified point are frequently denoted by x and y. The important fact for us at present is that if the coordinates of P are designated by x and y, then the radius vector which goes from 0 to P is given by $\rho = xi + yj$.

Cartesian coordinates in space. If to the pair of vectors $\boldsymbol{i}$ and j already introduced we add a third unit vector k perpendicular to the plane of i and j, then we have the basis for a space coordinate system which is a natural generalization of the plane coordinate system just considered. Specifically, the point corresponding to the number triplet x, y, z is the vertex of terminus of the radius vector ρ given by $\rho = xi + yj + zk$. Thus to locate the point correlated to 2, 3, 4, we first generate the vector $2i$ by moving two units in the direction in which i points, then we turn through a right angle and generate $3j$, and finally generate $4\boldsymbol{k}$ by moving 4 units in the $\boldsymbol{k}$-direction (which is perpendicular to the plane of i and j). The vertex of $4\boldsymbol{k}$ in the present position of $4\boldsymbol{k}$ is then the desired point. If i points from left to right, while j points away from us, and $\boldsymbol{k}$ points upward, then the vector ρ, $\rho = 2i + 3j + 4\boldsymbol{k}$ is generated by starting at the origin moving two units to the right, then three units away from us (j-direction), and finally four units straight upwards. If instead of positive numbers we had considered one or more negative numbers, the procedure would have been the same with the proviso, of course, that multiplication of a vector by a negative number calls for a reversal in direction.

A change of coordinate systems. The transformation equations connecting two rectangular Cartesian coordinate systems with a common origin is of fundamental importance for a variety of topics in mathematics and mathematical physics. As an illustration of the algebraic use of geometric entities, we shall derive the transformation equation connecting the base vectors i, j and $\bar{i}$, $\bar{j}$, of two given Cartesian coordinate systems in the plane. It is convenient to introduce the rotational operator i which turns vectors through a right angle in the positive (counter-clockwise) sense. We first recall our assumption that any vector in a plane may be represented in terms of a set of base vectors i, j. Hence, numbers l and m exist such that $\bar{i} = li + mj$. We now assume that the right angles between the coordinate vectors i, j and $\bar{i}$, $\bar{j}$, are both positive so that $i\bar{i} = \bar{j}$ while $ii = j$ and $ij = -i$. The rotational

operator i, has the effect of rotating the vector to which it is applied through an angle of 90° in the positive sense which we have taken to be the sense of the shortest rotation of i into j and hence of j into $-i$. From the natural assumption that the rotation of a triangle through 90° effects a rotation of all three sides through 90°, it follows that the operator i is distributive. Hence if we multiply both members of the equation $\bar{i} = li + mj$ by the rotational operator i, we get $\bar{j} = lii + mij = lj - mi$ since $i\bar{i} = \bar{j}$, $ii = j$, $ij = -i$. Thus the equations connecting the new base vectors $\bar{i}$, $\bar{j}$ with the original base vectors i, j are

$$\bar{i} = li + mj, \qquad \bar{j} = -mi + lj \tag{4.1}$$

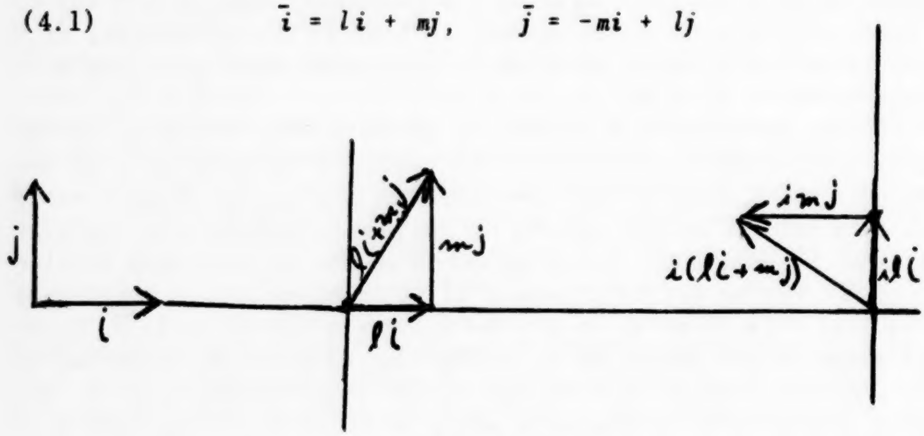


Fig. (4.2)

The square array of coefficients of i and j in the right members of (4.1), when the second equation is written directly beneath the first, namely, $\begin{pmatrix} l & m \\ -m & l \end{pmatrix}$ is called the *matrix* of the transformation. This particular matrix has some remarkable properties which persist through the generalization of (4.1) to three and higher dimensional cases. Transformations of this type are of sufficient importance in mathematics and mathematical physics to warrant a listing of these properties. First, since $\bar{i}$ is of unit length and equal to $li + mj$, it follows from the theorem of Pythagoras that $l^2 + m^2 = 1$ and this means that the sum of the squares of the elements in any row or column is *one*. Also, the sum of the products of corresponding elements in two different rows or in two different columns is *zero*. To illustrate, using the two rows (horizontal) l, m and $-m$, l, we have $l(-m) + ml = 0$. It will be very helpful for an understanding of later developments if we note carefully that this expression $l(-m) + ml$ is *the sum of the products of corresponding components of the vectors* $\bar{i}$ and $\bar{j}$ since $\bar{i} = li + mj$ and $\bar{j} = -mi + lj$. Furthermore, if we multiply the vector $\bar{i}$ by an arbitrary number a and $\bar{j}$ by some number b, the resulting vectors would still be at *right angles* and the *sum of the products of corresponding components*

would still be zero, thus $(al)(-bm) + (am)(bl) = 0$. If we abbreviate by denoting the components of $a\bar{i}$ by $a_1 a_2$ and similarly the components of $b\bar{j}$ by b_1, b_2, then we have $a_1 b_1 + a_2 b_2 = 0$. Evidently, the expression $a_1 b_1 + a_2 b_2$ is of some importance since its *vanishing* is apparently connected with the fact of *orthogonality*. Later on, we shall see that this certainly is the case.

In order to illustrate that the algebraic procedures involved in the manipulation of vector equations are no more difficult than those occurring in elementary algebra, let us attempt to solve the equations (4.1) for the vectors i and j in terms of the vectors $\bar{i}$, and $\bar{j}$. Examination shows that this can be done by the very same procedures used in the corresponding problem in elementary algebra. Thus multiplying the first equation by l and the second by $-m$ and adding, we get successively $l\bar{i} = l^2 i + lmj$, $-m\bar{j} = m^2 i - mlj$, $l\bar{i} - m\bar{j} = i$, since $l^2 + m^2 = 1$. Similarly, we have $j = m\bar{i} + l\bar{j}$ and, therefore, the complete set of equations connecting $\bar{i}$, $\bar{j}$ and i, j is

$$\bar{i} = li + mj, \quad \bar{j} = -mi + lj; \quad i = l\bar{i} - m\bar{j}, \quad j = m\bar{i} + l\bar{j}. \tag{4.2}$$

However, in one respect at least the vector equations are quite different from their real number counterparts. As an illustration, we shall find the transformation of coordinates induced by the transformation in the base vectors (4.2). Thus, let P be any point and let x, y and $\bar{x}, \bar{y}$ denote the coordinates of P relative to the base sets i, j and $\bar{i}, \bar{j}$. This means that if ρ is the radius vector $\overrightarrow{OP}$ from the origin to the point in question then, on the one hand, $\rho = xi + yj$, while for essentially the same reason (the meaning of Cartesian coordinates) $\rho = \bar{x}\bar{i} + \bar{y}\bar{j}$. That is

$$xi + yj = \bar{x}\bar{i} + \bar{y}\bar{j}. \tag{4.3}$$

Substitution for i and j from the last two equations of the set (4.2) yields the vector equation.

$$\bar{x}\bar{i} + \bar{y}\bar{j} = xl\bar{i} - xm\bar{j} + ym\bar{i} + yl\bar{j}. \tag{4.4}$$

We have now come to an instance of an important principle of vector algebra: *A vector equation in N-dimensional space is equivalent to N scalar* (or *real number*) *equations*. In the case of (4.4), N is two; and the essential fact is that any vector in the plane at hand has a unique set of components relative to $\bar{i}$ and $\bar{j}$. This means that the multiplier of $\bar{i}$ in the left member of (4.4) must equal the multiplier of $\bar{i}$ on the right. A similar statement holds of course with regard to $\bar{j}$. Hence from the *single* vector equation (4.4) there emerges the *pair* of scalar or real number equations,

$$\bar{x} = lx + my; \quad \bar{y} = -mx + ly \tag{4.5}$$

which constitute the expression of the new coordinates $\bar{x}$ and $\bar{y}$ in terms of the original coordinates x, y. Comparison of equations (4.5) with (4.1) reveals a remarkable fact; namely, the transformation of the base vectors (4.1) induces essentially the *same* (the technical term is *cogredient*) transformation in the coordinate variables, i.e. if in (4.1) we replace the vector symbols $\mathbf{i}$, $\mathbf{j}$, $\bar{\mathbf{i}}$, $\bar{\mathbf{j}}$, with the symbols for the coordinate variables x, y, $\bar{x}$, $\bar{y}$, then (4.1) becomes (4.5). Evidently, the solution of (4.5) for x and y is formally the same as the solution of (4.1) and hence we may read it out of the set of equations (4.1). Thus the complete set of coordinate transformation equations is

$$(4.6) \qquad \bar{x} = lx + my, \quad \bar{y} = -mx + ly; \quad x = l\bar{x} - m\bar{y}, \quad y = m\bar{x} + l\bar{y}.$$

Furthermore, since the *components* of a vector are merely the *coordinates* of its end point, when the vector is drawn from the origin, the components of a vector follow the transformation (4.6). To illustrate, if α is a vector in the plane at hand and a_1, a_2 are its components relative to the base set $\mathbf{i},\mathbf{j}$ while $\bar{a}_1$, $\bar{a}_2$ are its components relative to $\bar{\mathbf{i}},\bar{\mathbf{j}}$, so that $\alpha = a_1\mathbf{i} + a_2\mathbf{j} = \bar{a}_1\bar{\mathbf{i}} + \bar{a}_2\bar{\mathbf{j}}$, then

$$(4.7) \qquad \bar{a}_1 = la_1 + ma_2, \quad \bar{a}_2 = -ma_1 + la_2.$$

Similarly, if β is a second vector and b_1, b_2; $\bar{b}_1,\bar{b}_2$ are its components relative to $\mathbf{i}$, $\mathbf{j}$ and $\bar{\mathbf{i}}$, $\bar{\mathbf{j}}$, respectively, then

$$(4.8) \qquad \bar{b}_1 = lb_1 + mb_2, \quad \bar{b}_2 = -mb_1 + lb_2.$$

If we recall that the orthogonality of the vectors α and β implied the vanishing of the expression $a_1b_1 + a_2b_2$ and take account of the fact that a_1, a_2; b_1, b_2 are the components of α and β relative to *any* set of base vectors in the plane at hand, we see that whenever $a_1b_1 + a_2b_2$ vanishes, $\bar{a}_1\bar{b}_1 + \bar{a}_2\bar{b}_2$ must do so likewise. This remarkable fact leads us to investigate the effect of applying the transformation equations (4.7), (4.8) to the expression $\bar{a}_1\bar{b}_1 + \bar{a}_2\bar{b}_2$. The algebra is quite simple and so we shall carry out the work in detail, thus $\bar{a}_1\bar{b}_1 + \bar{a}_2\bar{b}_2 = (la_1 + ma_2)(lb_1 + mb_2) + (-ma_1 + la_2)(-mb_1 + lb_2) = (l^2 + m^2)a_1b_1 + 0a_1b_2 + 0a_2b_1 + (m^2 + l^2)a_2b_2 = a_1b_1 + a_2b_2$. We now have reached the remarkable conclusion that $\bar{a}_1\bar{b}_1 + \bar{a}_2\bar{b}_2 = a_1b_1 + a_2b_2$ or, in more technical language, we have discovered that the expression $a_1b_1 + a_2b_2$ is *invariant in form* under a rotation of base vectors. The background ideas involved here are as follows: Given a plane, a pair of unit-orthogonal vectors $\mathbf{i},\mathbf{j}$ and a vector α in the plane, there is a unique set of numbers a_1, a_2 such that $\alpha = a_1\mathbf{i} + a_2\mathbf{j}$. The numbers a_1, a_2 were named the components of α. If we draw a new set of base vectors $\bar{\mathbf{i}}$, $\bar{\mathbf{j}}$

(unit-orthogonal), then as before α can be represented in terms of $\bar{i}$, $\bar{j}$, thus, $\alpha = \bar{a}_1\bar{i} + \bar{a}_2\bar{j}$. Thus a change in the set of base vectors induces a change in the components of a vector. The actual relationship between the two sets of components in the case examined ($j = ii$, $\bar{j} = i\bar{i}$ rather than $\bar{j} = -i\bar{i}$) is given by equation (4.7). Since α is *any* vector in the plane, similar remarks apply to the representation of any second vector β. The sum of the products of corresponding components of two vectors α, β (namely, $a_1b_1 + a_2b_2$) however does not change but is *invariant*.

Interpretation of the invariant $a_1a_2 + b_1b_2$. Since the sum of the products of corresponding components of two given vectors α and β in the plane at hand has a value which does not change when we change the base vectors, we are free to simplify the computation by taking a special choice of base vectors. Obviously, if we take i along α then $\alpha = ai$ (a is the length of α) and therefore $a_1 = a$, $a_2 = 0$. Also, if (α,β) denotes the angle generated by the shortest rotation of α into β, then by definition the cosine and sine of (α,β) are merely b_1/b and b_2/b if b_1, b_2, and b denote respectively, the coordinates of the end point of β when β is drawn from the origin, and the length of β. The coordinates b_1, b_2 are also the *components* b_1, b_2 of β, thus the *components* of β are given by the equations $b_1 = b\cos(\alpha,\beta)$, $b_2 = b\sin(\alpha,\beta)$. Consequently, we have $a_1b_1 + a_2b_2 = ab\cos(\alpha,\beta) + 0\cdot b\sin(\alpha,\beta) = ab\cos(\alpha,\beta)$.

5. **The dot product.** The main features of the so called dot product of two vectors α and β are readily deducible from the extended relationship

$$\alpha\cdot\beta = ab\cos(\alpha,\beta) = a_1b_1 + a_2b_2. \tag{5.1}$$

Here $\alpha\cdot\beta$ (read α dot β) is a new symbol defined by the first relationship. The numbers a_1, a_2; b_1, b_2 as in the preceding section are the components of α and β relative to a set of base vectors i, j in a plane containing α and β. By assumption a vector (unrestricted) can be displaced from one point to another without change. Now since $\alpha\cdot\beta$ by definition is $ab\cos(\alpha,\beta)$, to know the meaning of the dot product it is sufficient to understand the symbols appearing in the product $ab\cos(\alpha,\beta)$. Obviously, the dot product is a process which produces a single *real number* from a pair of vectors α and β, but there are of course infinitely many such processes (there are infinitely many functions of the angle (α,β) for example) and so two pertinent questions arise at once. Why emphasize this particular process by dignifying it with a special name and why call it a product? With regard to the latter query, it is customary to call a process which produces a unique entity from a given pair of entities x and y a product if it is distributive; i.e., if $x(y + z) = xy + xz$, this of course implies that the operation, or operations, + occurring here have already been defined. Hence the central question with regard to the propriety of calling $\alpha\cdot\beta$ a product is this: Are the two real numbers $\alpha\cdot(\beta + \gamma)$ and $\alpha\cdot\beta + \alpha\cdot\gamma$ equal? The

key case is that in which γ lies in the α,β-plane. One method of attack is to prove the desired relationship by means of a figure. An alternative procedure is to use the equality $\alpha\cdot\beta = a_1b_1 + a_2b_2$ and attack the problem algebraically. We shall adopt the latter procedure and denote the components of γ by c_1 and c_2 so that $\beta + \gamma = b_1i + b_2j + c_1i + c_2j = (b_1 + c_1)i + (b_2 + c_2)j$. Thus from the fact that the dot product of two vectors is equal to the sum of the products of their corresponding components, we have

$$\alpha\cdot(\beta+\gamma) = a_1(b_1+c_1) + a_2(b_2+c_2) = a_1b_1 + a_2b_2 + a_1c_1 + a_2c_2 = \alpha\cdot\beta + \alpha\cdot\gamma.$$

As a matter of fact, the fortunate circumstance that $\alpha\cdot\beta$ is expressible as the sum of the simple products a_1b_1 and a_2b_2 assures us at once that it $(\alpha\cdot\beta)$ has many of the properties of real numbers, and points the way to easy but important generalizations involving more dimensions. For example, since the order of the factors in the multiplication of real numbers is immaterial, we have $\alpha\cdot\beta = a_1b_1 + a_2b_2 = b_1a_1 + b_2a_2 = \beta\cdot\alpha$. Similarly, since the multiplication of real numbers is associative and distributive we have $x(a_1b_1 + a_2b_2) = (xa_1)b_1 + (xa_2)b_2 = a_1(xb_1) + a_2(xb_2)$ and therefore $x(\alpha\cdot\beta) = (x\alpha)\cdot\beta = \alpha\cdot(x\beta)$.

On the basis of this evidence it is obviously justifiable to call the dot process a *product*. There remains, however, the question of its relative importance. For the present, perhaps it will suffice to say that it provides a very powerful method for the investigation of certain matters of a geometrical or physical nature and that the natural generalizations of the expression $a_1b_1 + a_2b_2$ play a conspicuous part in various higher dimensional geometries and physical theories. As an indication of the applicability of this product to elementary geometry and physics, we note that if α is a unit vector, then $\alpha\cdot\beta$ is equal to the signed projection of β on to α. On the other hand, if α represents the displacement of a particle while acted on by a constant force β, $\alpha\cdot\beta$ is the work done on the particle by the force during the displacement. *Projection* and *work* are both concepts of cardinal importance. Perhaps it should be emphasized at this point that the vanishing of the dot product means that either one or both of the vectors vanish or that the vectors are at right angles.

Before considering a specific application of the dot product, it will be helpful to observe that because the distributive and commutative laws hold, the products of sums and differences of vectors may be expanded as in ordinary algebra. To illustrate:

$$(5.2)\qquad (\alpha - \beta)\cdot(\alpha - \beta) = (\alpha - \beta)\cdot\alpha + (\alpha - \beta)\cdot\beta =$$

$$\alpha\cdot(\alpha - \beta) + \beta\cdot(\alpha - \beta) = a^2 - 2ab\cos(\alpha,\beta) + b^2.$$

Here we have used the fact that $\alpha\cdot\alpha = a^2$ and $\beta\cdot\beta = b^2$ (since a and b

denote the magnitudes of α and β), and $\cos 0 = 1$).

The law of cosines. An insight into the power, simplicity, and naturalness of vector methods may be obtained by discovering and proving the celebrated *law of cosines*, one of the most important theorems of elementary mathematics. The problem is to discover and establish a relationship between the lengths of the sides of a triangle. As a first step in the investigation, we draw a triangle and convert the sides into vectors by introducing arrowheads. We let two of the vectors say α and β issue from a common point so that the third vector γ will be equal to their difference. We denote the magnitudes of these vectors by a, b, c and denote by the corresponding capitals the angles opposite these sides. Since $\gamma\cdot\gamma = c^2$ equation (5.2) may be rewritten as $c^2 = a^2 + b^2 - 2ab \cos C$, which is the desired relation.

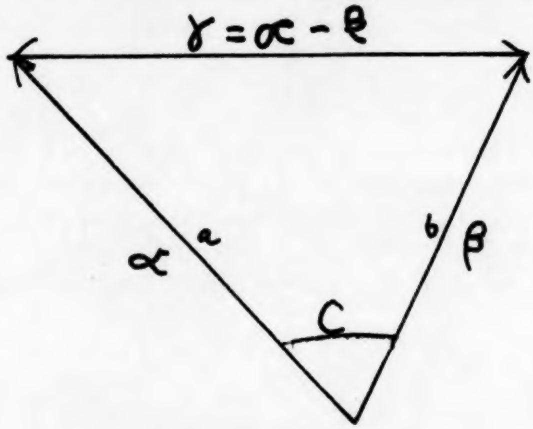


Fig. (5.1)

6. **The equations of certain simple loci.** Since a vector or directed stroke is itself a piece of a line, it is to be expected that the simplest and most intuitive equations of straight lines are the vector equations. As a first step in developing these equations we introduce the radius vector ρ. This particular vector, it will be recalled, has its origin or initial point anchored at a certain given point O which will usually be taken as the origin of coordinates if a coordinate system is subsequently introduced. It is required in general that ρ be made to rotate and vary in magnitude in such a way that its vertex or terminal point will trace out the locus in question. The central problem in developing the vector equation of a locus is to find the conditions that ρ must satisfy in order that its end point may pick up all of the points of the locus and no others.

Line through a given point in a given direction. As a first illustration, we present the equation of the line through the origin and having the direction of a given vector β. Evidently, this equation is

simply $\rho = u\beta$ with the proviso that u be a real variable of unrestricted magnitude. It will be recalled that for u positive, $u\beta$ is a vector having the direction of β but possessing a magnitude equal to u times the magnitude of β; while for u negative, the direction of $u\beta$ is opposite to that of β. Thus we see that the role of the multiplier u is to stretch and contract and in addition to effect a reversal whenever it is negative. Thus since the origin of ρ is confined to the point O, the equation $\rho = u\beta$ restricts ρ to the infinite collection of vectors obtainable by drawing β from O and then stretching and contracting β and its reversal. Thus for $u = 1$, ρ is the vector β and the point determined is the end point of β. The point determined by $u = 2$ is the end point of 2β, while $u = \frac{1}{2}$ is correlated to the midpoint of β, β being drawn from the the origin. Similarly, $u = -1$ yields the end point of the reversal of β.

In a like manner the equation of the line through a given point A and parallel to a given vector β is simply

$$\rho = \alpha + u\beta \tag{6.1}$$

if α is the fixed radius vector $\vec{OA}$. The idea here is simply that all points on the given line and only these points may be obtained by stepping from the origin to A and then stepping various distances in the direction of β and $-\beta$. Here α serves to take us from the origin up to the line, while β acts as a pointer to indicate the direction of the line.

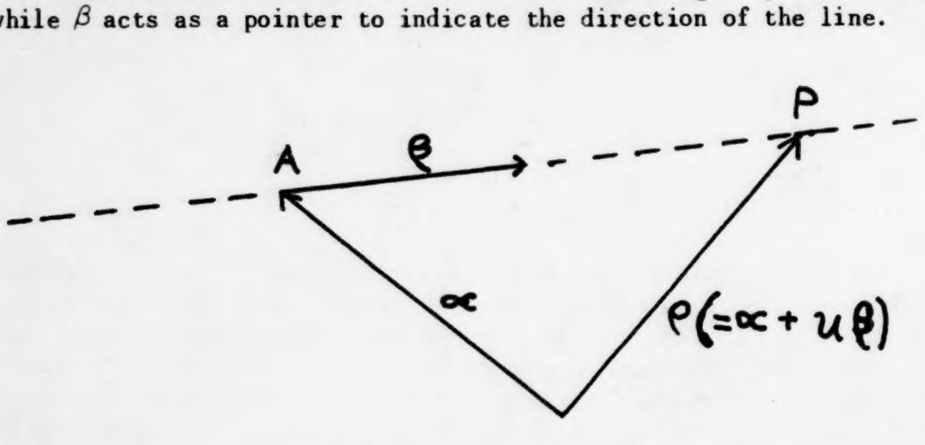


Fig. (6.1)

The normal equation of a plane. As a final application of the dot product, we develop the vector equation of the plane through the end of a given radius vector δ and perpendicular to δ. As before, we let ρ be the variable radius vector whose vertex generates the locus in question—in this case the plane. Now since ρ and δ both terminate in the plane, $\rho - \delta$ lies in the plane and is therefore perpendicular to δ, thus we have at once

(12.7) $$(\rho - \delta) \cdot \delta = 0.$$

To visualize the geometrical background for this equation we may take for the given plane a table top and let δ be one of the legs so that the origin is on the floor, then since every line on the table top is perpendicular to the leg we see that $\rho - \delta$ which lies on the table top must be perpendicular to δ. The vector ρ is a variable vector which begins at the bottom of δ and ends in the table top.

7. **To more general theories and higher dimensionality.** We have assumed that relative to a given set of base vectors, each vector has a unique set of components and conversely that each set of components determines a unique vector. The introduction of the components in a sense achieves an arithmetization of the geometrical concept vector. Consequently, we can think and work either in terms of a vector as a geometric entity (a directed line segment or stroke) or in terms of its arithmetical representation - the collection of its components.

Components versus vectors. For some purposes (such as the problem of deriving the equation of a line through a given point and parallel to a given vector), it seems to the writer to be better to work directly with the vectors themselves as geometric entities rather than with their components. On the other hand, certain key expressions in the components such as $a_1b_1 + a_2b_2$ suggest very clearly how we may create certain higher dimensional geometries that will constitute satisfying generalizations of the familiar two and three dimensional Euclidean geometries.

As a matter of fact, a variety of fruitful questions emerge from consideration of the components such as: Is it possible satisfactorily to attach a meaning to the term components of a vector relative to a curvilinear coordinate system (i.e., one in which the curves obtained by holding all of the coordinates but one fixed and allowing that one to vary are not straight lines in all cases)? Assuming this problem to be solved satisfactorily, what are the properties of the transformation equations connecting the components in two different coordinate systems of the more general type? Can systematic procedures be devised for the construction of invariants relative to this wider group of coordinate transformations? Can a satisfactory two dimensional vector analysis be built up out of the collection of vectors tangent to a surface such as the surface of a sphere? Such questions may lead to extensive mathematical developments and perhaps shed some light on the way mathematics grows.

Multiple algebra. If in the two dimensional case, we center our attention on the two components (a_1, a_2) of a given vector α it is a natural step, in retrospect at least, to attempt the formulation of an algebra of number pairs. The facts relating to the components of vectors suggest the following definitions $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$; $x(a_1, a_2) = (xa_1, xa_2)$. The first of these interpreted vectorially means that the components of $\alpha + \beta$ are $a_1 + b_1$, $a_2 + b_2$ - the sums of the corresponding components. With regard to multiplication,

from our point of view only distributive processes are admissible as products. If we define the product of two number pairs (a_1, a_2) and (b_1, b_2) by the equation $(a_1, a_2)(b_1, b_2) = a_1b_1 + a_2b_2$, then the theory of course duplicates our two dimensional vector algebra. It should be observed, however, that multiple algebra has other possibilities. For example, the complex number $a_1 + ia_2$ can be represented by the number pair (a_1, a_2). Since $(a_1 + ia_2) + (b_1 + ib_2)$ is defined to be $a_1 + b_1 + i(a_2 + b_2)$, we have as before $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$. However, a new product $(a_1, a_2)(b_1, b_2)$ must now be introduced defined as the number pair $(a_1b_1 - a_2b_2, a_1b_2 + a_2b_1)$ to correspond to the definition $(a_1 + ia_2)(b_1 + ib_2) = a_1b_1 - a_2b_2 + i(a_1b_2 + a_2b_1)$. The number pairs (a, o) and (o, b) are of course identified with the real number a and the pure imaginary ib. The originally musterious number $\sqrt{-1}$ now appears as the simple number pair $(0,1)$ and $(0,1)(0,1) = (-1,0) = -1$. Complex numbers are frequently represented by vectors in the complex plane. The multiple algebra viewpoint brings out the similarities and differences between these two vector algebras—that of Gibbs and that of complex numbers. On the algebraic side both of the parent theories of Gibbs' vector analysis, the *theory of quaternions* and the *Ausdehnungslehre* are multiple algebras.

Higher dimensionality. Our next concern is to examine the problem of extending the notion dot product to three and higher dimensions. To make the step to three dimensions, we assume as in the section on Cartesian coordinates in space that we have given three mutually orthogonal unit vectors which we now denote by $\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$ rather than by three different letters. By assumption, a given vector α has a unique representation of the form $\alpha = a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + a_3\mathbf{i}_3$. The numbers a_1, a_2, a_3 are, of course, called the somponents of α relative to the given base vectors. Thus associated with each vector α there is a number triple (a_1, a_2, a_3). If we proceed from the standpoint of multiple algebra, we are led as in the two-dimensional case to the definitions $(a_1, a_2, a_3) + (b_1, b_2\ b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$ and $x(a_1, a_2, a_3) = (xa_1, xa_2, xa_3)$. Obviously, the most direct generalization of the double algebra equivalent of the dot product, namely, the product defined by $(a_1, a_2)(b_1, b_2) = a_1b_1 + a_2b_2$ is the product defined by the equation $(a_1, a_2, a_3)(b_1, b_2, b_3) = a_1b_1 + a_2b_2 + a_3b_3$. If we take this to be the expression for $\alpha\cdot\beta$ in terms of the components relative to $\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$, then $\alpha\cdot\beta$ will be distributive and $\alpha\cdot\alpha$ will reduce to $(a_1)^2 + (a_2)^2 + (a_3)^2$, the sum of the squares of the components. This last expression is the square of the length of α in Euclidean geometry since α, $(\alpha = a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + a_3\mathbf{i}_3)$ is a diagonal of a rectangular box of sides a_1, a_2, a_3. Incidentally, if we consider successively the vector $a_1\mathbf{i}_1$, then the vector $a_1\mathbf{i}_1 + a_2\mathbf{i}_2$ obtained by adding $a_2\mathbf{i}_2$ to the vertex of $a_1\mathbf{i}_1$, and finally the vector $a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + a_3\mathbf{i}_3$ obtained by adding $a_3\mathbf{i}_3$ to the preceding vector, we see that the high hurdle in computing their magnitudes occurs in the two-dimensionzl case. The length squared of the first vector is simply

$(a_1)^2$, while the length of the second vector is the length of the hypotenuse of a right triangle, and hence the theorem of Pythagoras is needed to show that $|a_1 i_1 + a_2 i_2|^2 = (a_1)^2 + (a_2)^2$. However, we don't need a new theorem to prove that $|a_1 i_1 + a_2 i_2 + a_3 i_3|^2 = (a_1)^2 + (a_2)^2 + (a_3)^2$. The reason is that the added vector $a_3 i_3$ is perpendicular to every vector in the (i_1, i_2)-plane, and therefore $(a_1 i_1 + a_2 i_2)$, taken as a single vector, and $a_3 i_3$ form a right triangle. The square of the length of one leg is $(a_1)^2 + (a_2)^2$, while the square of the length of the other is $(a_3)^2$, consequently, the square of the hypotenuse is $(a_1)^2 + (a_2)^2 + (a_3)^2$. If we take the next step and postulate the existence of a fourth vector i_4 which is perpendicular to each of i_1, i_2, i_3 and to every vector expressible in terms of them, then the vector $a_4 i_4$ should be regarded as perpendicular to $(a_1 i_1 + a_2 i_2 + a_3 i_3)$ considered as a single vector and it would seem that the theorem of Pythagoras is again applicable and if so it would give $|a_1 i_1 + a_2 i_2 + a_3 i_3 + a_4 i_4|^2 = (a_1)^2 + (a_2)^2 + (a_3)^2 + (a_4)^2$. This of course is a matter of pure supposition. In special relativity there is a consistent four dimensional geometry and the key expression may be written in the form $(a_1)^2 + (a_2)^2 + (a_3)^2 - (a_4)^2$. Thus there is not just one four dimensional geometry but many such. It is of course necessary to build up a higher dimensional geometry carefully from a self-consistent postulational basis. In the writer's opinion, it is better to proceed algebraically rather than by means of situations existing in two and three dimensional geometry. One might inadvertently assume, say, that if two planes have a point in common then they have at least a line in common. It isn't necessary to investigate higher dimensional geometry in detail to feel uneasy about this assumption provided we look in the right direction. To illustrate let us set $\rho = xi_1 + yi_2 + zi_3 + wi_4$, define the coordinate planes in the usual way, and then examine the vector equations for the (x,y)- and (z,w)-planes, namely

$$\rho = xi_1 + yi_2 + 0i_3 + 0i_4, \quad \rho = 0i_1 + 0i_2 + zi_3 + wi_4.$$

Equating corresponding components, we get as the one and only common point, the origin $x = 0$, $y = 0$, $z = 0$, $w = 0$. Again, if we let x, y, z be the coordinates in ordinary Euclidean space and let t denote the time, we may construct as in the theory of relativity a four dimensional manifold by taking *point* to mean the abstraction correlated to x, y, z, t – a location at a definite instant. If we set $z = 0$, $t = 0$, with x and y variable, we get the (x,y)-plane of our four-space, which from the three-dimensional viewpoint is the ordinary x,y-plane at the instant $t = 0$ only. The (z,t)-plane consists of the set of all points on the

z-axis, not for just one particular value of t, but for all values of t. The common ground of the (x, y)- and (z, t)-planes is again the "origin" $x = 0$, $y = 0$, $z = 0$, $t = 0$.

Granting the need for careful procedure in the construction of higher dimensional geometries, it is urgent that such a concept as say *orthogonality* unless it is selected as a basic undefined term should be defined explicitly. Orthogonality is more complicated in three dimensions than in two. The algebraic approach provides an easy way to precise definition. One way to proceed is to adopt a generalization of the dot product by generalizing the key expression $a_1b_1 + a_2b_2$ and then define $|\alpha|$, orthogonality of α and β, $\cos(\alpha, \beta)$ through the generalization adopted, thus: $|\alpha| \equiv \sqrt{\alpha\cdot\alpha}$, orthogonality of α and β means $|\alpha| \neq 0$, $|\beta| \neq 0$, $\alpha\cdot\beta = 0$; $\cos(\alpha, \beta) \equiv \alpha\cdot\beta/(|\alpha|\,|\beta|)$, for $\alpha \neq 0$, $\beta \neq 0$.

The generalization, suggested by multiple algebra, of $a_1b_1 + a_2b_2$ is easily verified in the case of ordinary Euclidean 3-space with a set of base vectors i_1, i_2, i_3. Since $\alpha - \beta = (a_1 - b_1)i_1 + (a_2 - b_2)i_2 + (a_3 - b_3)i_3$ it follows that $|\alpha - \beta|^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2 = (a_1)^2 + (a_2)^2 + (a_3)^2 + (b_1)^2 + (b_2)^2 + (b_3)^2 - 2(a_1b_1 + a_2b_2 + a_3b_3) = |\alpha|^2 + |\alpha|^2 - 2(a_1b_1 + a_2b_2 + a_3b_3)$. But since the vectors in the dot product $(\alpha - \beta)\cdot(\alpha - \beta)$ can all be contained in a single plane, we have $|\alpha - \beta|^2 = (\alpha - \beta)\cdot(\alpha - \beta) = \alpha\cdot\alpha + \beta\cdot\beta - 2\alpha\cdot\beta = |\alpha|^2 + |\beta|^2 - 2\alpha\cdot\beta$. Comparison of these two expressions gives the desired relationship.

$$\alpha\cdot\beta = a_1b_1 + a_2b_2 + a_3b_3.$$

8. **The cross product.** Just as the vanishing of the dot product implies orthogonality, there is a second product whose vanishing expresses parallelism—it being supposed, of course, that the vectors themselves are not of zero length. The parallelism of two vectors α and β beans that $\alpha = k\beta$ or $a_1 = kb_1$, $a_2 = kb_2$, $a_3 = kb_3$ or $a_1/b_1 = a_2/b_2 = a_3/b_3$. By cross multiplying we get the following three equations $a_2b_3 - a_3b_2 = 0$, $a_3b_1 - a_1b_3 = 0$, $a_1b_2 - a_2b_1 = 0$. The left members of these equations are differences of products of components of α and β and, furthermore, there are *three* of them—just the number of components of a vector in three dimensional space. Thus we are led to consider the following vector expression:

$$(8.1)\quad i_1(a_2b_3 - a_3b_2) + i_2(a_3b_1 - a_1b_3) + i_3(a_1b_2 - a_2b_1).$$

We should note here that the order of the subscripts in the positive terms is in agreement with the cyclic order 1, 2, 3, 1, 2, thus 1, 2, 3; 2, 3, 1; 3, 1, 2. If we replace the components of β in the vector expression (8.1) with the components of $\beta + \gamma$, i.e., if we replace b_1 with $b_1 + c_1$, b_2 with $b_2 + c_2$, b_3 with $b_3 + c_3$, as we shall do presently,

it will become apparent that this vector expression which is constructed out of the components of the two given vectors has a *distributive property*. Hence (8.1) can be regarded as a product of the two vectors α and β. If the base set is right-handed (this means that i_3 has the direction of the thumb of the right hand when the curled fingers indicate the shortest rotation of i_1 into i_2) then it turns out that (8.1) is the so called *cross product* of α and β. We shall denote the *cross product* of α and β by the symbol $\alpha \times \beta$ and abbreviate (8.1) by just writing the first term $i(a_2 b_3 - a_3 b_2) + \cdots$. The usual definition of the cross product, which may be shown to be equivalent to that above, is as follows:

Definition (8.1) $\qquad \alpha \times \beta = ab|\sin(\alpha, \beta)|\nu.$

Here a and b denote the magnitudes of α and β and ν is a unit vector perpendicular to α and β and so directed that α, β, ν will constitute a right-handed system. If α and β are parallel or zero and hence do not determine ν, the product is assigned the value zero.

Carrying out the details of the check for the distributive property, we have

$$\alpha \times (\beta + \gamma) = i_1[a_2(b_3 + c_3) - a_3(b_2 + c_2)] + \cdots$$

$$= i_1(a_2 b_3 - a_3 b_2) + \cdots + i_1(a_2 c_3 - a_3 c_2) + \cdots = \alpha \times \beta + \alpha \times \gamma.$$

Unlike the dot product the cross product is *not* commutative but anti-commutative, this means that $\alpha \times \beta = -\beta \times \alpha$. This fact can be verified easily from the definition or by interchanging the order of the letters a and b in the expression (8.1). Also the cross product doesn't admit of easy generalization to higher order spaces - drastic revision is necessary. Like the dot product it is associative with respect to multiplication by a scalar s in the sense that

$$s(\alpha \times \beta) = (s\alpha) \times \beta = \alpha \times (s\beta)$$

The cross product finds wide and effective application in geometry and classical physics (particularly in mechanics and electricity and magnetism). Such fundamental concepts as the moments of force and momentum about a point 0 have their simplest expression in terms of the cross product.

The higher products. Of the products involving more than two factors the most fundamental are the *scalar triple product* $\alpha \cdot (\beta \times \gamma)$ and the *vector triple product* $\alpha \times (\beta \times \gamma)$. The scalar triple product has the following properties: (a) it is a scalar, (b) it is unaltered by an interchange of dot and cross (the vectors adjacent to the cross are to be grouped together), (c) its absolute value is equal to the volume of the parallelopiped determined by the vectors, (d) its vanishing means that one of the vectors is expressible in terms of the other two. This

product is fundamental in the step from unit orthogonal base vectors to the more general sets whose members are not orthogonal and not necessarily of equal length. If α, β, γ is such a set and $\alpha\cdot(\beta \times \gamma) \neq 0$, then any vector ρ is expressible in the form $\rho = x\alpha + y\beta + z\gamma$. To solve for x, y, z (the components of ρ relative to α, β, γ and likewise the coordinates of its end point in the new coordinate system when drawn from the origin), we multiply (dot product) by $\alpha^*\cdot$, $\beta^*\cdot$, $\gamma^*\cdot$, these new vectors being defined as follows: $\alpha^* = \beta \times \gamma/s$, $\beta^* = \gamma \times \alpha/s$, $\gamma^* = \alpha \times \beta/s$, $s = \alpha\cdot(\beta \times \gamma)$. The values of x, y, z turn out to be $\rho\cdot\alpha^*$, $\rho\cdot\beta^*$, $\rho\cdot\gamma^*$ respectively.

The outstanding property of the vector triple product is its expansion formula, namely, $\alpha \times (\beta \times \gamma) = (\alpha\cdot\gamma)\beta - (\alpha\cdot\beta)\gamma$. Thus $\alpha \times (\beta \times \gamma)$ is a vector expressible in terms of β and γ.

9. **The calculus of vectors.** It is impossible to give an adequate account of the differential and integral calculus of vectors without assuming an extensive knowledge of the calculus of real variables; however, a few brief remarks should suffice to give an idea of the flavor of this part of vector theory. The derivative of a vector function $\rho(t)$ with respect to t turns out to be a piece of the tangent line to the curve $\rho = \rho(t)$. The definition of the *derivative* follows the pattern of the definition of dy/dx, i.e., $d\rho/dt = \lim \Delta\rho/\Delta t$, $\Delta\rho = \rho(t_0 + \Delta t) - \rho(t_0)$. To illustrate, $\rho = \cos t\,i + \sin t\,j$ is the vector equation of a circle and here ρ is a function of t. The derivative $d\rho/dt$ is $-\sin t\,i + \cos t\,j$, a vector perpendicular to the radius vector $\cos t\,i + \sin t\,j$ (since $\rho\cdot d\rho/dt = 0$) and therefore tangent to the circle. If the radius vector ρ is expressed as a function of arc length s along the curve, then $d\rho/ds$ is a unit tangent vector, $d^2\rho/ds^2$ is perpendicular to the curve in what is called the principal normal direction, and $|d^2\rho/ds^2|$ is the curvature. Vector methods are ideal for the exposition of much of the differential geometry of curves and surfaces as well as for the discussion of mechanics, electromagnetism and other branches of physics.

The University of Texas

EXTERIOR BALLISTICS

John W. Green

Two of the greatest English mathematicians of all times are two who have been associated over the past three or four decades, co-authoring over a hundred mathematical papers and setting a standard for mathematical analysis so severe that there almost seems to be none worthy of following in their footsteps. These are G. H. Hardy and J. E. Littlewood.

Now Hardy is said to have once said of Littlewood, of whose technical facility he had the greatest respect, "Littlewood tried to make a decent science out of ballistics, and if Littlewood could not, who could?"

There is unfortunately a certain amount of truth suggested by Hardy's remark. I do not claim that ballistics is an indecent science, but it has to be admitted that it is an unbeautiful one, compared to many other branches of mathematics. It is applied mathematics at its extreme. One is not allowed to select problems to work on, but they are forced upon one; one does not follow up those aspects of a problem which seem attractive, but those which lead to an answer. There are a few exceptions to this - such as Gronwall's interesting work on the qualitative properties of the trajectory, but these are largely out of the domain of ballistics and into that of fancy. So far as I know, no bit of new basic mathematical knowledge has come from ballistics. However, into it vast quantities of knowledge obtained and techniques developed in other branches of mathematics have been poured.

Neither do I claim that ballistics is uninteresting. It can be, and frequently is interesting; however it is usually the subject matter of the physical problem which is the interesting thing, not the mathematics. There is a certain amount of glamour and morbid fascination to the subject.

Now up until fairly recently, most of the principal contributions to ballistics have come, not from professional mathematicians but from ordnance officers and astronomers. (An exception is Bashforth, an early British ballistician and clergyman!) The interest of the former in ballistics is obvious, and practically all the work up to 1914 is due to French, German, and Italian officers. These made many clever advances in the subject, and the theory developed by them was quite adequate for the weapons used before World War I.

The appearance of the astronomers on the scene date to 1914 or thereabouts, and is a manifestation of the phenomenon of total mobilization of resources which war demands these days. The astronomers have four important qualifications which make them successful ballisticians. First, they have adequate backgrounds in the laws of physics and mechanics to handle the physical problems involved. Second, they are adept at numerical processes and are capable and energetic computers. Third, they have developed in astronomy techniques of pertubation calculations that lend

themselves ideally to ballistic problems. Finally, and perhaps most important, they are able to construct and operate the fine optical and other equipment which is essential to ballistic instrumentation. As a result, in recent times of emergency, the astronomers have turned almost en masse to ballistics.

During World War I, numerous mathematicians also turned to ballistics and made important contributions, principally in numerical methods of trajectory computing. More and more mathematicians have been turning to the subject in recent years, particularly as the distinction between ballistics and aerodynamics gradually breaks down and introduces the mathematically more difficult problems of aerodynamics. At the Aberdeen Proving Ground, the Army's principal ballistic establishment, there was in World War II possibly the largest collection of mathematicians ever collected for an appreciable time.

Now what does one study in ballistics anyway? I should say that I am restricting myself to exterior ballistics, which deals with the behavior of the projectile from gun muzzle to striking point. Interior ballistics, which deals with what happens in the gun barrel, and terminal ballistics, which deals with the damaging possibilities on striking, are quite different matters.

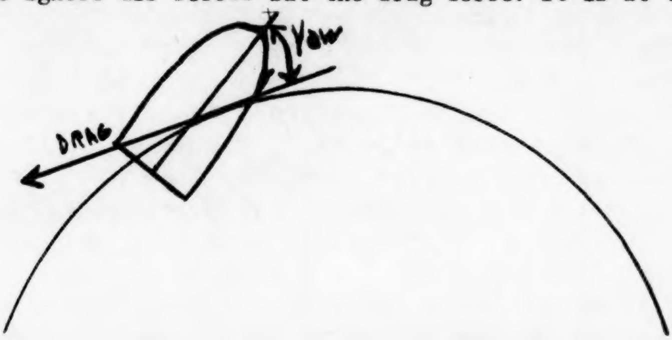
Exterior ballistics (before rockets, at least) consists of the study of the motion of an unpropelled projectile under the action of whatever forces act on it. These forces are the force of gravity and the aerodynamical forces of the atmosphere through which the projectile passes. The former is, of course, easily determined. It is proportional to the mass of the projectile and toward the center of the earth. Right here is a good place to introduce one of the most important ballistic facts of life - namely that there is no ballistic problem approximating a physical situation which is sufficiently simple to solve exactly. The keywords in ballistics are, simplify - approximate - idealize. So almost always the force of gravity is assumed to be constant in magnitude and direction instead of obeying Newton's beautiful law. Not that one cannot solve dynamic problems with the exact gravitational law; one can, but taken in conjunction with the other complications usually present in ballistic problems, this refinement costs too much and is usually ignored. But more of this slightly later.

Now when we turn to the second kind of force the matter is quite different. The motion of the atmosphere in the neighborhood of a projectile is very complicated, and to discover what it is, one has *very* difficult partial differential equations to solve. Yet if one intends to find by theory and not by experiment the forces on a projectile, one must solve the equations. No one* has *ever* found by this method the forces on *any* projectile which would agree with those experimentally determined well enough to allow their use. What is done? One measures the force experimentally. This may be done in many ways, the use of

*With the possible exceptions of the sphere and cone (infinite).

wind tunnels being the most obvious. That is, we hang the projectile in a wind tunnel and measure the force on it as a function of wind velocity.

Now the system of force is very complicated. The principal force is called the drag and is directed opposite to the direction of motion. However the projectile does not always point in the direction it is going, and so a transverse force, or lift, results. If the projectile is spinning, other forces act. However, what is usually done as a first approximation is to ignore all forces but the drag force. It is at this



stage of approximation that ballistics proper sets in. That is, the study of the motion of a projectile is actually a difficult aerodynamic problem. However, if all the forces are determined experimentally and synthesized into a single system of force depending on the velocity of the projectile, the resulting equations are ordinary differential equations instead of partial, and the study is called ballistics.

Now the equations above have been the cause of much back-breaking labor and the matter of the discovery of many clever ways of approximate solution—particularly by the early Italian and French ballisticians. They are not usually susceptible of solution in terms of elementary functions, and the modern method consists mainly of the strong-arm method of numerical step-by-step integration. There have been discovered three ways of effectively carrying this out by devices and machines of one sort or another. The first is the young lady named Mildred, or the like, who computes on the desk computer and grinds out a trajectory in perhaps 12-48 hours. The second is the differential analyzer, which does it in a few seconds. In spite of these figures, the first is one of the best.

Now when the trajectory is computed, one knows where the projectile will land. It is, however, one of the bitterest disappointments to those studying ballistics that this is not the way it is usually done. The military is decidedly averse to accepting the ballisticians' prediction as to where the projectile will land. (Aren't mathematicians notoriously careless with decimal points and such things?) In fact, the military is disgustingly pragmatic in these matters and has always insisted on shooting the gun and measuring the range to see where the projectile goes. If this is the case, where does the ballistician fit in, if at all? The answers lie principally in the fact that it is not feasible to make

a sufficient number of experiments to determine the behavior of the projectile under all conditions where it might be used. Actually what is done is usually this: A few firings are made, the drag function estimated from the results, and the other trajectories computed on the basis of this drag function.

Also, of course, in the case of bombs dropped from airplanes where the initial data are so hard to determine, and in anti-aircraft firing, where the final data are difficult to determine experimentally, the ballistician becomes even more essential.

One of the most important problems which the mathematical ballistician has to solve is that of effects of small changes in conditions of one sort or another. Take for example the effect of the non-constancy of gravity and the rotation of the earth. The effect of these and many other similar non-standard conditions can be computed with the use of suitable mathematical analysis of certain differential equations related to the equation of the trajectory. Now when would one use, say, a correction for the rotation of the earth? This correction is not seriously large except on the very longest trajectories, and is very complicated to apply, depending as it does on the angle of fire, the initial velocity, the type of projectile, the latitude, and the direction of fire. Where it is used is in the reduction of the range firings which are carried out to make the firing table. A 40-foot error due to earth rotation may not be a bad thing to tolerate, particularly if the error is random due to firing in different directions. But if this were ignored in the range firings, it would prejudice *every* trajectory by this 40 feet and always in the same direction.*

Now up until the last war, the most complicated problems which the ballistician had to solve were those connected with the spinning shell—and this is no simple problem. It is treated like a particle, plus correction for the fact that it isn't. The gyroscopic effect of the spin has to be taken into account. But in the last war, new and more complicated problems arose. The first of these I mention are those in connection with the flight of bombs. Now a bomb trajectory is really just the last half of an artillery trajectory, so that there is nothing new there. However, the initial velocities are very small (400 feet/sec.), and the method of launching is such as frequently to introduce large yaws. (The yaw is the angle between the direction the projectile is going and the direction it is pointing.) Besides, bombs are usually light for their size, compared to shells, since they consist of mostly explosive and little steel. As a result, the force system is far from reducing to a simple drag. There is a considerable lift force also. Furthermore, the yaw does not remain in the same direction; it oscillates back and forth about the zero position. The result is a rather complicated analysis

*Who says the direction is random? Not in a given war, anyway.

of differential equations with oscillatory solution, and an increase in the excitement in the life of the ballistician.

The second kinds of new problem I mention are those in connection with rockets. Imagine everything as before, but with a force of variable strength propelling the rocket in the direction in which it is pointed. A very complicated problem, particularly since some of them spin also. These new problems give rise to new and more interesting mathematical problems on which ballisticians are spending their time.

There is another place in ballistics where mathematicians and astronomers especially have been employed, and this is in the making of photographs of projectile in flight and measuring these photographs and reducing these measurements to physical data on position, velocity, and acceleration. The mathematics used has no appreciable connection with the theory of exterior ballistics, and indeed consists mainly of difficult problems in solid geometry and trigonometry and computing. However, where there is computing, there is mathematics; namely the mathematics that provide ways of reducing the amount of computing.

I have tried to give an idea of some of the problems which ballisticians encounter. Most of the people engaged in ballistics find their work nearly as much engineering as mathematics. In my approximate three years as a practicing ballistician, I found myself doing jobs as diverse as computing a trajectory on a desk computing machine, organizing fairly large scale trajectory computing projects for other computers, preparing bombing tables, designing a bomb-sight cam to build the firing table into the sight, and designing a fitting by which to hang a 1,000-pound rocket to swing so that its moment of inertia could be found.

Unfortunately ballistics seems to be here to stay, and employment possibilities in it are very good at the present. I hope we all live to see the day when it will exist no longer at all as a branch at least of *applied* mathematics. By the time of that millenium, perhaps someone more able than Littlewood will convert it into a respectable member of the family of branches of pure mathematics.

University of California, Los Angeles

GAMES OF STRATEGY

Melvin Dresher

Introduction. We shall review briefly a relatively new branch of mathematics - the theory of games of strategy. This includes parlor games as chess, bridge, poker; economic games as the production and pricing of goods, and military games as defense against attack and timing of attack. In such games the players can influence the outcome of the game by outwitting each other. We shall omit games of chance, such as roulette, whose outcome depends entirely on chance and cannot be influenced by the actions of the players except by cheating.

Games of chance have been studied mathematically for many years, and the mathematical theory of probability was developed from their study. However, the problems of games of strategy were hardly considered until 1928 when von Neumann wrote the first mathematical paper on the subject. This paper was later expanded jointly with Morgenstern in their book, *Theory of Games and Economic Behavior, published in* 1945.

Von Neumann defined a game very generally. It is a set of rules which describe what moves each player may or must make under all possible circumstances, what moves are made for the players by chance, what information is available to the players and what payments are made to the players when the game is over.

Initially, each game is described in terms of its moves, which may be personal (the moves are made by the decision of the players), or chance (the moves are made for the players by some random device). The first simplification of the description of a game is the introduction of the concept of a strategy. In the actual play of a game, each player, instead of making his decision at each move of the game, may formulate a complete plan for playing the game from beginning to end, covering all possible contingencies. Such a plan is referred to as a *strategy*. The plan or strategy includes any information that may become available to the player in accordance with the rules of the game. In this way no freedom of action is lost through the use of a strategy since a strategy specifies the player's action in terms of the information that might become available. In effect, a player selects a strategy and the strategy plays the game.

Each possible way that a player has for playing a game from beginning to end is a strategy for that player. If we enumerate all his possible ways of playing the game, we shall obtain all the strategies for that player. Every combination of strategies, one from each player, will determine an outcome of the game which is measured by the payoff to the player. The totality of payoffs to a player can be represented by a *payoff matrix*.

The fundamental problem of the theory of games of strategy is to determine for each player his best strategies, or the best ways to

play a given game from beginning to end. We shall discuss only games having two players and where the winnings of one player represents the losses of the other. All parlor games are of this type.

Two Person "le Her". We can illustrate the concepts of the previous section by means of the 18th century card game "le Her". Such famous mathematicians as N. Bernoulli and Montmort attempted to solve this game but failed. Although the game was played by several persons, we shall describe it for two players. A dealer gives, at random, one card to a receiver and one to himself. The object of the game is to obtain the higher card. If the receiver is dissatisfied with his card he may compel the dealer to change with him, but if the dealer has a king, the dealer is allowed to retain it. If the dealer is dissatisfied with the card he holds, after the receiver has used his option, the dealer may exchange it for another card drawn at random from the deck. However, if the card he then draws is a king, he must keep the card with which he was dissatisfied. If the dealer and receiver have equal cards, the dealer wins.

This game consists of three moves. The first move is a chance move - dealing one card each, at random, to the receiver and dealer. The second move is a personal move by the receiver - exchanging his card with the dealer or staying with the original card. The third move is a personal move by the dealer - exchanging his card with a card from the deck or staying with the card he holds.

The game can be summarized by defining a strategy for the receiver to be a determination of change or stay for each of the 13 cards. One such strategy might be written as

1	2	3	4	5	6	7	8	9	10	J	Q	K
c	*s*	*c*	*s*	*s*	*c*	*c*	*s*	*s*	*c*	*s*	*c*	*s*

where c means change and s means stay. This strategy tells the receiver to change if he receives an ace, stay if he receives a two, change if he receives a three, ..., change if he receives a queen, stay if he receives a king. Note that the strategy is a complete set of instructions. It is apparent that the receiver has 2^{13} strategies. The dealer's strategy can be described in the same way, and therefore he also has 2^{13} strategies. It is not necessary to compute each payoff in the entire $2^{13} \times 2^{13}$ matrix, for it is apparent that any strategy in which exactly r changes of cards appear is dominated by a strategy in which the first r cards are changes and the remaining $13 - r$ cards are holds. It can be verified that of the 2^{13} strategies of each player, $2^{13} - 2$ of them are poor strategies - two strategies dominate all the others. Thus after all the poor strategies are removed, the game of "le Her" reduces to the following 2×2 payoff matrix

PAYOFF TO DEALER

(Probability of Winning)

		Receiver's Strategies	
		Hold 7 and over	Change 7 and under
Dealer's Strategies	Hold 8 and over	16,182 / 33,150	16,122 / 33,150
	Change 8 and under	16,146 / 33,150	16,182 / 33,150

We shall solve this game in a later section.

Minimax Theory. The mathematical model of the two-person zero-sum game can be described in the following way: Player I chooses a strategy i from his m strategies and player II chooses a strategy j from his n strategies, each choice being made without any knowledge of the other. The payoff to player I is a_{ij} and the payoff to player II is $-a_{ij}$. Player I wishes to maximize a_{ij} but he controls only the choice of i; player II wishes to maximize $-a_{ij}$, or to minimize a_{ij} but he controls only the choice of j. Any solution of the game should resolve these conflicting interests.

Player I has some strategy, i^*, such that he can receive at least $\max_i \min_j a_{ij}$. Player II has some strategy, j^*, such that he pays at most $\min_j \max_i a_{ij}$. In general,

$$a_{ij^*} \leq \min_j \max_i a_{ij} \geq \max_i \min_j a_{ij} \leq a_{i^*j}$$

for all i and j. If the game is such that $\max_i \min_j a_{ij} = \min_j \max_i a_{ij}$, then we refer to i^*, j^* as *optimal* strategies of player I and II respectively, and $a_{i^*j^*}$ as the *value* of the game. The optimal strategies have the property that if player I chooses i^*, he is certain to receive at least $a_{i^*j^*}$, and if player II uses j^* he can make sure that $a_{i^*j^*}$ is the most that I receives. Even if player I were to announce in advance that he plans to play strategy i^*, player II could not take advantage of this information and thereby reduce I's winnings. A necessary and sufficient condition that a game have optimal strategies i^*, j^* is that there exists an element $a_{i^*j^*}$ of the payoff matrix which is simultaneously the minimum of its row and the maximum of its column.

If the payoff matrix $A = \|a_{ij}\|$ is such that

$$\max_i \min_j a_{ij} < \min_j \max_i a_{ij},$$

then there exists no single strategy that is optimal. In such a game a player can pick a best strategy only if he knows the strategy chosen by his opponent. But by the rules of the game this is impossible for the players must pick their strategies simultaneously. In such a game it is also important that a player's chosen strategy be unknown to his opponent. This can be accomplished by mixing the strategies. A player, instead of choosing a single strategy, chooses a probability distribution over the whole set of strategies and a random device selects a particular strategy, subject to the given probability distribution. Such a probability distribution is called a *mixed strategy*.

The fundamental theorem of games of strategy, first proven by von Neumann, states that if each player has a finite number of strategies, then each player has an optimal mixed strategy. That is, player I has a mixed strategy such that if he uses it he can expect to win at least an amount v, and player II has a mixed strategy such that if he uses it he can expect to lose no more than the amount v. Thus there exist mixed strategies X_0, Y_0, for players I and II respectively such that

$$X'AY_0 \leq v = X_0'AY_0 \leq X_0AY,$$

for all mixed strategies X and Y of the two players. The optimal mixed strategy also has the property that it may be announced, in advance, and an opponent cannot take advantage of this information.

Solution of "le Her". Applying the fundamental theorem to the game "le Her" , we find that the optimal strategy for the dealer is to mix his two strategies with probabilities $(\frac{3}{8}, \frac{5}{8})$, respectively, and for the receiver it is to mix his strategies with probabilities $(\frac{5}{8}, \frac{3}{8})$. The value of the game to the dealer is .487 and to the receiver it is .513. It is interesting to note that the dealer is at a disadvantage in this game, although all the rules seem to be in his favor. This is partially explained by the fact that the receiver has the first move and is in a more powerful position to influence the outcome of the game.

Three-Finger "Morra". A well-known finger game player by two people is the game of morra. Each player shows one, two, or three fingers and simultaneously guesses the number of fingers his opponent will show. If just one player guesses correctly, he wins an amount equal to the sum of the fingers shown by himself and his opponent; otherwise the game is a draw. The payoff matrix is given by:

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
(1,1)	0	2	2	-3	0	0	-4	0	0
(1,2)	-2	0	0	0	3	3	-4	0	0
(1,3)	-2	0	0	-3	0	0	0	4	4
(2,1)	3	0	3	0	-4	0	0	-5	0
(2,2)	0	-3	0	4	0	4	0	-5	0
(2,3)	0	-3	0	0	-4	0	5	0	5
(3,1)	4	4	0	0	0	-5	0	0	-6
(3,2)	0	0	-4	5	5	0	0	0	-6
(3,3)	0	0	-4	0	0	-5	6	6	0

This game has an infinite number of optimal mixed strategies. Every convex linear combination of the following four mixed strategies is an optimal strategy:

$$\left(0,\ 0,\ \frac{5}{12},\ 0,\ \frac{4}{12},\ 0,\ \frac{3}{12},\ 0,\ 0\right)$$

$$\left(0,\ 0,\ \frac{16}{37},\ 0,\ \frac{12}{37},\ 0,\ \frac{9}{37},\ 0,\ 0\right)$$

$$\left(0,\ 0,\ \frac{20}{47},\ 0,\ \frac{15}{47},\ 0,\ \frac{12}{47},\ 0,\ 0\right)$$

$$\left(0,\ 0,\ \frac{25}{61},\ 0,\ \frac{20}{61},\ 0,\ \frac{16}{61},\ 0,\ 0\right)$$

Since the game is symmetric, the value of the game is zero.

Games with Infinite Numbers of Strategies. So far we have assumed that each player has a finite number of strategies. Many games have an infinite number of strategies. For example, economic games in which the strategy consists of determination of a price or a percentage, have a continum of strategies. The mathematical model for such a game is the following: Player I chooses a strategy x and player II chooses a strategy y, each satisfying $0 \leq x \leq 1$, $0 \leq y \leq 1$. The payoff to player I is given by a function $M(x,y)$. Since the game is zero-sum, the payoff to player II is $-M(x,y)$.

Whereas every finite game has an optimal strategy, some infinite games do not have optimal strategies. However, if $M(x,y)$ is a continuous function of x and y, then each player always has an optimal strategy. In general it is very difficult to determine the optimal strategies of an infinite game. We give some examples of infinite games and their solution.

Allocation of Forces. Two targets of value l and k are to be defended against attack. How should the defender distribute his total defensive

force among the two targets and how should the offense distribute his attacking force among the two targets? A convenient payoff to the offense can be defined by

$$M(x,y) = \begin{array}{ll} x - y & \text{if } x \geq y \\ k(x - y) & \text{if } x \leq y \end{array}$$

where x is the fraction of the offensive force sent against the target of value 1, and y is the fraction of the defensive force placed at that target.

Solution: The offense mixes his strategy - he attacks a single target with all his forces but selects the target at random subject to the following distribution function

$$F(x) = \frac{1}{1+k} I_1(x) \frac{k}{1+k} I_k(x),$$

where $I_a(x)$ is a step-function with unit jump at $x = a$.

The defense uses a pure strategy - places $\frac{1}{1+k}$ of his force at target with value 1, and $\frac{k}{1+k}$ of his force at target with value k.

Expectation: Offense can expect to win $\frac{k}{1+k}$.

Games of Timing. The theory of games of strategy can be used to solve a large class of timing problems. In these problems, the rules of the game describe the actions which the players are to take, but the timing of the actions is left to the players. The situation is such that each player wishes to delay the actions as long as possible but he is also penalized for delaying. This conflict of interests can be resolved and a best timing of action can be determined for each player. A duel between two persons is an example of a game of timing.

Game of Timing with Information. Each duelist is informed of his opponent's action, firing, as soon as it takes place. The two duelists close in on each other with no opportunity for retreat. The accuracies increase steadily. The payoff is 1 to the surviving duelist and 0 if both or neither survives. In terms of the accuracies, the payoff is given by the following payoff function:

$$M(x,y) = \begin{array}{ll} 2P_1(x) - 1 & \text{if } x < y \\ P_1(x) - P_2(x) & \text{if } x = y \\ 1 - 2P_2(y) & \text{if } x > y. \end{array}$$

where $P_1(x)$, $P_2(y)$ are the accuracies of the two duelists at time x and y respectively.

Solution: Each duelist fires at x^* where $P_1(x^*) + P_2(x^*) = 1$, if his

opponent has not fired already. If opponent has fired at x^*, then fire at x such that $P_1(x) = 1$.

Expectation: Player I wins $P_1(x^*) - P_2(x^*)$.

Game of Timing without Information. Each duelist is ignorant of any firing by the other. Assume that the duelists have the same accuracy. functions. Then the payoff function is given by:

$$M(x,y) = \begin{cases} -y + (1+y)x & \text{if } x < y \\ 0 & \text{if } x = y \\ -y + (1+y)x & \text{if } x > y \end{cases}$$

where x and y are the kill probabilities.

Solution: Never fire before $x = \frac{1}{3}$. From $x = \frac{1}{3}$ to $x = 1$, fire with density

$$dF(x) = \frac{1}{4x^3}\,dx.$$

REFERENCES

1. von Neumann - Morgenstern, "Theory of Games and Economic Behavior," Princeton (1945).
2. Todhunter, "A History of the Mathematical Theory of Probability," (1865), pp. 106-110.
3. "Contributions to Theory of Games," *Annals of Mathematics, Study* 24, Princeton (1950).
4. J. McDonald, "A Theory of Strategy," *Fortune*, June 1949, pp. 100-110.
5. E. W. Paxson, "Recent Developments in the Mathematical Theory of Games," *Econometrica* 17 (1949), pp. 72-73.
6. M. Dresher, "Methods of Solution in Game Theory," *Econometrica* 18 (1950), pp. 179-180.

The Rand Corporation

MISCELLANEOUS NOTES

Edited by

Charles K. Robbins

Articles intended for this Department should be sent to Charles K. Robbins, Department of Mathematics, Purdue University, Lafayette, Indiana.

THE LOG LOG SCALES OF THE SLIDE RULE

In certain computations using the log log scales on a slide rule, there may be uncertainty either about the section in which the answer is located or about the position of the decimal point. A procedure for eliminating these doubts is presented herewith. Pertinent to this procedure are the following definitions:

Definition 1. When the slide extends to the right at the completion of an operation, such condition is designated as *Status I*.

Definition 2. When the slide extends to the left at the completion of an operation, *Status II* exists.

Definition 3. The *adjusted characteristic* of the logarithm of a number is simply the characteristic plus unity. *Adjustment* is made when Status II exists.

The procedure mentioned above is based upon the following rule. This rule is expressed with the understanding that the sections *LL*1, *LL*2, and *LL*3 are numbered 1, 2, and 3, respectively; and, likewise, *LL*01, *LL*02, and *LL*03 are also numbered 1, 2, and 3, respectively.

Rule. *Consider $Z = X^Y$. The number of the section on which a power Z is found is equal to the number of the section on which the base X is found plus the characteristic of the logarithm of the exponent Y (the latter being adjusted if required by status).*

Proof: To facilitate the presentation of the proof, the following nomenclature is adopted:

k_j = the number of the section bearing the number J;

C_j = the characteristic of the logarithm of J;

M_j = the mantissa of the logarithm of J;

$\overline{C}_j$ = the characteristic of the logarithm of J, adjusted when required by the *ststus* of an operation.

In terms of these symbols, we wish to prove that for $Z = X^Y$,

$$k_z = k_x + \overline{C}_y$$

With reference to $Z = X^Y$, suppose X to be greater than unity. This is the case where the log log scales greater than one ($LL1$, $LL2$, and $LL3$ on some slide rules) are used. Taking the natural logarithm of both members, we obtain:

$$\ln Z = (\ln X)Y$$

Now taking the common logarithm of both sides, we have:

$$\log(\ln Z) = \log(\ln X) + \log Y$$

or, $$M_{\ln z} + C_{\ln z} = M_{\ln x} + C_{\ln x} + M_y + C_y \qquad (1)$$

If there is no carryover upon adding $M_{\ln x}$ and M_y (Status I exists), then

$$M_{\ln z} = M_{\ln x} + M_y \qquad (2)$$

Substituting (2) in (1):

$$C_{\ln z} = C_{\ln x} + C_y \qquad (3)$$

If there is a carryover upon adding $M_{\ln x}$ and M_y (Status II exists), then

$$M_{\ln z} + 1 = M_{\ln x} + M_y \qquad (4)$$

Substituting (4) in (1):

$$C_{\ln z} = C_{\ln x} + (C_y + 1) \qquad (5)$$

We may summarize (3) and (5) in the expression:

$$C_{\ln z} = C_{\ln x} + \overline{C}_y \qquad (6)$$

where $\overline{C}_y$ is the characteristic of the logarithm of Y, adjusted when required by the status of the operation.

We will now show that a result similar to (6) holds for the case, $X < 1$. If X in $Z = X^Y$ is less than unity, then $\ln Z$ and $\ln X$ in

$$\ln Z = (\ln X)Y$$

are both negative. Hence, it is necessary to multiply both members by -1, before taking the common logarithm of both sides:

$$\log(-\ln Z) = \log(-\ln X) + \log Y$$

Or, $$C_{-\ln z} + M_{-\ln z} = C_{-\ln x} + M_{-\ln x} + C_y + M_y$$

From this, following the process used for $X > 1$, we derive:

$$C_{-\ln z} = C_{-\ln x} + \overline{C}_y \tag{7}$$

Now consider any number P within range of the log log scales. If we take $P > 1$, we would find it on a section of the log log scales greater than one. If $P < 1$, we would find it on a section of the new log log scales less than one (*LL*01, *LL*02, and *LL*03 on some slide rules). By consulting a table of natural logarithms, we can show that if

P is greater than one, $\quad C_{\ln p} = k_p - 3;$

P is less than one, $\quad C_{-\ln p} = k_p - 3.$

Since P was any number, for X and Z, in particular, we have:

$$C_{\ln x} = k_x - 3 \tag{8}$$

$$C_{\ln z} = k_z - 3 \tag{9}$$

and,
$$C_{-\ln x} = k_x - 3 \tag{10}$$

$$C_{-\ln z} = k_z - 3 \tag{11}$$

Substituting (8) and (9) in (6), and substituting (10) and (11) in (7), we obtain a common result, namely:

$$k_z = k_x + C_y$$

Now note the following examples:

Example 1. Evaluate 6.35^3.
Solution: Set the hairline to 6.35 which is on the *LL*3-scale. Draw the left *C*-index under the hairline. Move the hairline to 300 on the *C*-scale. Since Status I exists, $k = 3 + 0 = 3$. Thus, the answer, 256, is read on the *LL*3-scale.[1]

Example 2. Evaluate $2.335^{0.75}$.
Solution: Set the hairline to 2.335 which is on the *LL*2-scale. Draw the right *C*-index under the hairline. Move the hairline to 750 on the *C*-scale. Since Status II exists, $k = 2 + (-1 + 1) = 2$. Thus, the answer, 1.889, is read on the *LL*2-scale.

Example 3. Evaluate $275^{0.004}$.
Solution: Set the hairline to 275 on the *LL*3-scale. Draw the right

1. It may be helpful to note that in taking the natural logarithm of both members of $6.35^3 = x$: $(\ln 6.35)3 = \ln x$, we have immediate indication that a multiplication is to be performed in which 6.35 is to be located on a log log scale, 3 on the *C*-scale, and in which x is to be read from a log log scale.

C-index under the hairline. Move the hairline to 400 on the C-scale. Since Status II exists, $k = 3 + (-3 + 1) = 1$. Thus, the answer, 1.0227, is read on the $LL1$-scale.

Example 4. Evaluate $0.05^{0.0612}$.
Solution: Set the hairline to 0.05 on the new log log scales less than one, or on the $LL03$ scale. Draw the right C-index under the hairline. Move the hairline to 612 on the C-scale. Since Status II exists, $k = 3 + (-2 + 1) = 2$. Thus, the ansser, 0.832, is read on the proper $LL2$ or $LL02$ scale under the hairline.

Example 5. Find x in $17 = x^4$.
Solution: We note that $x = 17^{1/4}$. Set the hairline to 17 on the $LL3$-scale. Draw 400 on the C-scale under the hairline. Since division is used and since Status II exists, $k = 3 - (0 + 1) = 2$. Thus, $x = 2.03$ since the left C-index is over 2.03 on $LL2$.

Example 6. Evaluate $7.31^{26/41.5}$.
Solution: Set the hairline to 7.31 on the $LL3$-scale. First draw 415 on the C-scale under the hairline. At this stage, Status II exists—therefore, $k = 3 - (1 + 1) = 1$. Move the hairline to 260 on the C-scale. Status II still exists, so $k = 1 + (1 + 1) = 3$. Thus, the answer, 3.48, is read on the $LL3$-scale.

Note: The problems which follow are of the type, $Z = X^Y$, where X and Z are known constants and Y ix to be found. Since Y is a value which is to be taken from the C-scale, the position of the decimal point must be determined. In accordance with the rule given in this article, $k_z = k_x + \overline{C}_y$; or $\overline{C}_y = k_z - k_x$. But, C_y equals either C_y or $C_y + 1$ according as Status I or Status II exists, respectively. So, when Status I exists $C_y = k_z - k_x$; and when Status II exists, $C_y = k_z - k_x - 1$.

Example 7. Find x in $15 = 3^x$
Solution: Set the hairline to 3 on $LL3$. Draw the left C-index under the hairline. Move the hairline to 15 on $LL3$. Since Status I exists, $C_x = 3 - 3 = 0$. Thus, the answer, 2.46, is on C under the hairline.

Example 8. Find x in $9^x = 1.72$.
Solution: Set the hairline to 9 on $LL3$. Draw the left C-index under the hairline. Move the hairline to 1.71 on $LL2$. Since Status I exists, $C_x = 2 - 3 = -1$. Thus, the ansser, 0.247, is on C under the hairline.

Example 9. Find x in $2.03^x = 17$.
Solution: Set the hairline to 2.03 on $LL2$. Draw the right C-index under the hairline. Move the hairline to 17 on $LL3$. Since Status II exists, $C_x = 3 - 2 - 1 = 0$. Thus, the answer, 4, is on C under the hairline.

University of Tennessee E. L. Eagle

CURRENT PAPERS AND BOOKS

Edited by

H. V. Craig

This department will present comments on papers previously published in the MATHEMATICS MAGAZINE, lists of new books, and book reviews.

In order that errors may be corrected, results extended, and interesting aspects further illuminated, comments on published papers in all departments are invited.

Communications intended for this department should be sent in duplicate to H. V. Craig, Dept. of Applied Mathematics, Univ. of Texas, Austin 12, Texas.

The Theory of Algebraic Numbers. By Harry Pollard, Carus Mathematical Monograph, No. 9. Math. Assoc. of America, 1950, 142 pages.

This monograph gives a concise account of the elementary parts of nineteenth century algebraic number theory, culminating in applications to Fermat's last theorem and a proof of the Dirichlet-Minkowski unit theorem. The exposition is clear, and the proofs elegantly condensed.

The hypothetical reader for whom the book is intended is not only assumed to be ignorant of German (otherwise he would read Hecke and Hilbert) but also of the most elementary facts of ring theory as set forth for example in the first three chapters of McCoy's Carus Monograph *Rings and Ideals*. On the other hand, a certain mathematical maturity is required to follow the details of the proofs. Perhaps a senior or first-year graduate student, docile and intelligent, with an excellent grasp of calculus, but only a rudimentary acquaintance with modern algebra and modern languages, best meets these requirements.

For other readers, there are the great German classics, or the admirable Princeton Monograph of Weyl on Algebraic Numbers. This monograph is certainly more difficult than Pollard's. In compensation, the reader who masters it will understand not only Dedekind's ideas, but also the root ideas of Kronecker, Hensel and Hilbert. It is these ideas which have led to the development of modern valuation theory and class field theory, the main research fields for twentieth century work in algebraic numbers.

Morgan Ward

The Arithmetic Theory of Quadratic Forms. By Burton W. Jones, Carus Mathematical Monograph, No. 10, Math. Assoc. of America, 1950, 212 pages.

This excellent monograph gives a complete and self-contained account of the central ideas of the arithmetic theory of quadratic forms. Only the fundamentals of matrix algebra and elementary number theory are presupposed, but the reader is led by easy stages to the very frontiers of present-day knowledge.

The arithmetic theory of quadratic forms has been completely trans-

formed in the last thirty years by the researches of Hasse, Siegel, Pall and Jones, himself. In fact many of the methods of proof and results are the author's own, and are presented here for the first time. Certainly the book should do much to encourage further research in this fascinating field.

The book is beautifully planned and written, distinguished not only by what is included, but by what has been left out. It is certainly one of the best of the Carus Monographs published to date, and no one at all interested in the Theory of Numbers can afford to miss it.

Morgan Ward

Editor's Note

The editors of the Mathematics Magazine like either to accept submitted articles or give good and sufficient reasons for not accepting them.

However papers on classic problems that have been proved unsolvable, such as trisecting all angles and squaring the circle, have become so numerous that it is not feasible for us to unravel them. Hence we have adopted the policy that authors will be required to show us some error in the classic proofs that such problems are unsolvable before we will weigh their attempts to solve them.

PROBLEMS AND QUESTIONS

Edited by

C. W. Trigg, Los Angeles City College

Readers of this department are invited to submit for solution problems believed to be new and subject-matter questions that may arise in study, in study, in research, or in extra-academic situations. Proposals should be accompanied by solutions, when available, and by such information as will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted.

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and twice the size desired for reproduction. Readers are invited to offer heuristic discussions in addition to formal solutions.

Send all communications for this department to *C. W. Trigg, Los Angeles City College, 855 N. Vermont Ave., Los Angeles 29, California.*

PROPOSALS

112. *Proposed by Victor Thébault, Tennie, Sarthe, France.*

Find a number of the form *aaabbbccc* which gives, when increased by unity, a perfect square of nine digits.

113. *Proposed by Benjamin Greenberg, Brooklyn, N. Y.*

Isosceles triangles with base angles of 30° are constructed externally on the sides of triangle *ABC*. The third vertices of the isosceles triangles determine an equilateral triangle. Can this be proven by pure synthetic geometry without recourse to trigonometry?

114. *Proposed by Dewey Duncan, East Los Angeles Junior College.*

a) In air a gold surfaced sphere weighs 7588 grams. It is known that it may contain one or more of the metals aluminum, copper, silver or lead. When weighed successively under standard conditions in water, benzene, alcohol, and glycerine its respective weights are 6588, 6688, 6778, and 6328 grams. How much, if any, of the forenamed metals does it contain, if the specific gravities of the designated substances are taken to be as follows?

Aluminum	2.7	Alcohol	0.81
Copper	8.9	Benzene	0.90
Gold	19.3	Glycerine	1.26
Lead	11.3	Water	1.00
Silver	10.5		

b) If the ball is given a very thin coating of plastic, to prevent amalgamating, how far will it sink when floating in mercury, specific gravity 13.6?

115. *Proposed by Leo Moser, Texas Technological College.*

If all the faces of a polyhedron have central symmetry, prove that at least 6 of the faces are parallelograms. (A parallelopiped has exactly 6 such faces.)

116. *Proposed by H. H. Berry, University of Kentucky.*

If a, b, and c are integers such that $a + b = c$ and the least common multiple of a and b is M, where $ab/(a,b) = M$, then $(c,M) = (a,b)$. [The symbol (a,b) represents the greatest common divisor of a and b.]

117. *Proposed by H. D. Grossman, New York, N.Y.*

If n is a positive integer, prove that: a) The maximum difference between any two of the four sums, $\binom{n}{0} + \binom{n}{4} + \binom{n}{8} + \cdots$, $\binom{n}{1} + \binom{n}{5} + \binom{n}{9} + \cdots$, $\cdots$, $\binom{n}{3} + \binom{n}{7} + \binom{n}{11} + \cdots$ is $2^{[n/2]}$, where $[n/2]$ means the greatest integer $\leq n/2$.
β) The maximum difference between any two of the six sums, $\binom{n}{0} + \binom{n}{6} + \binom{n}{12} + \cdots$, $\binom{n}{1} + \binom{n}{7} + \binom{n}{13} + \cdots$, $\cdots$, $\binom{n}{5} + \binom{n}{11} + \binom{n}{17} + \cdots$, is $2 \cdot 3^{(n/2)-1}$ if n is even, or $3^{(n-1)/2}$ if n is odd.

118. *Proposed by A. M. McKinney, Dallas, Texas.*

Can the following integrals be evaluated? a) $\int e^x \sec x\, dx$; b) $\int e^{ix} \sec x\, dx$, without using the relation $e^{ix} = \cos x + i \sin x$; c) $\int e^{-ix} \sec x\, dx$, without using the relation $e^{-ix} = \cos x - i \sin x$.

SOLUTIONS

Late Solution

85. *Erich Michalup, Caracas, Venezuela.*

Parallel Tangents to a Quartic

89. [Jan. 1951] *Proposed by H. T. R. Aude, Colgate University.*

If the graph of the quartic $y = f(x) = x^4 + px^2 + qx + s$ has points of inflection, then there exist uniquely three pairs of parallel lines which are tangent to the quartic. Find the equations of the six lines, and note that the sum of their three slopes is $3q$.

Solution by Alan Wayne, Flushing, N.Y. Clearly "points of inflection" refers to "*real* points of inflection" which exist if and only if p is negative. Also, we note that it is desired to show that there exist uniquely three sets of precisely two parallel lines tangent

to the quartic.

This problem has been solved, at least implicitly, by J. S. Frame [Tangent triangles to a biquadratic curve, *American Mathematical Monthly*, **51**, 445-450, (1944)]. We make use of Frame's results and apply a simple calculation. By means of the "vertical affine" transformation

$$X = (-6/p)^{1/2}x, \qquad Y = (36/p^2)(y - qx - s) + 9, \tag{1}$$

the given quartic is reduced to the "normal form"

$$Y = F(X) = X^4 - 6X^2 + 9, \tag{2}$$

with preservation of collinearity of points, concurrency or parallelism of lines, midpoints, centroids, tangents and contact points, and ratios of distances along parallel lines.

Let T_1, T_2 and T_3 denote respectively the bitangent, the left inflectional tangent, and the right inflectional tangent of the normal quartic defined by (2), and let U_i, $(i = 1, 2, 3)$, be the tangent to the quartic which is parallel to T_i. Then (T_i, U_i), as Frame shows, are the only three sets in which there afe two (instead of three) parallel tangents. These lines have the following equations:

$$\begin{array}{llll} T_1 : & Y = 0. & U_1 : & Y = 9 \\ T_2 : & Y = 8X + 12. & U_2 : & Y = 8X - 15 \\ T_3 : & Y = -8X + 12. & U_3 : & Y = -8X - 15 \end{array}$$

Denote the images of these lines, respectively, by t_i and u_i, under the transformation (1) which takes the quartic from normal form back into the "general" form. The desired equations are then

$$m_1 = q, \quad m_2 = q + 8(-p/6)^{3/2}, \quad m_3 = q - 8(-p/6)^{3/2},$$

$$b_1 = s - (p^2/4), \quad b_2 = b_3 = s + (p^2/12),$$

$$c_1 = s \quad \text{and} \quad c_2 = c_3 = s - (2p^2/3).$$

Thus $m_1 + m_2 + m_3 = 3q$.

Also solved by *the proposer*.

For other properties of the graph of the quartic polynomial see *American Mathematical Monthly*, **37**, 510, (1930); **56**, 106, 165-170, (1949); *National Mathematics Magazine*, **15**, 429, (May 1941); and *This MAGAZINE*, **22**, 73-76, (Nov.-Dec. 1948).

A Square Covered by Three Unit Squares

93. [March 1951] *Proposed by Leo Moser, Texas Technological College.*

What is the size of the largest square that can be completely covered by three unit squares?

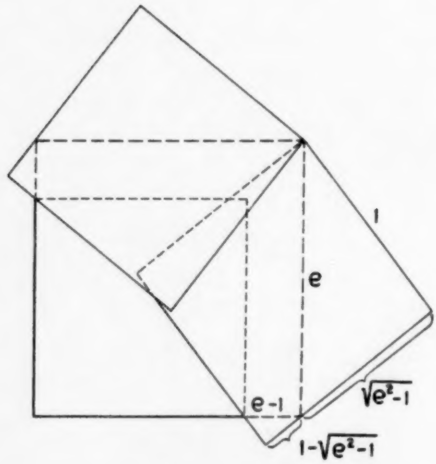


Figure 1

Solution by Michael Goldberg, Washington, D.C. Place one small square in a corner of the large square. Place the other two small squares so that a vertex of each is hinged at the opposite vertex of the large square, overlapping each other as shown in Figure 1. These small squares should lie so that their edges pass through the other vertices of the large square and the nearby vertices of the small square in the corner.

Let the edge of the large square be represented by e. Then the lengths of other line segments are functions of e as represented in Figure 1. From the similarity of the two right triangles, we have

$$e/1 = (e - 1)/(1 - \sqrt{e^2 - 1}).$$

Hence

$$e^4 - e^2 - 1 = 0 \quad \text{or} \quad e = \sqrt{(1 \quad \sqrt{5})/2} \doteq 1.272.$$

Also solved by *Leon Bankhoff, Los Angeles, California; and T. C. Wilderman, Baltimore, Md.*

Bankhoff also offered a solution for the case where overlapping is not permitted. He placed two unit squares side by side and rested the third square symmetrically on the other two as in Figure 2. He then found the diagonal of the largest square that can be inscribed within the three-square configuration. Assuming that the sides of the desired square pass through A and B, the locus of one end (P) of the diagonal will be the semicircle AGB (radius = 1/2) on AB. The locus of the other end (Q) of the diagonal will be the line CE. Since the diagonal bisects angle APB it passes through F the extremity of the diameter FG perpendicular to AB. Let x be the length of the diagonal and represent angle PFG by θ. Then

$$x = PF + FG = \cos\theta + (\sec\theta)/2.$$

Setting $dx/d\theta = 0$, we have $\sec\theta\tan\theta - 2\sin\theta = 0$, from which

$$\cos\theta = 1/\sqrt{2}, \qquad \theta = 45°, \qquad x = \sqrt{2} \text{ (Minimum diagonal)}$$

or

$$\sin\theta = 0, \qquad \theta = 0, \qquad x = 3/2 \quad \text{(Maximum diagonal).}$$

Consequently the side of the maximum square which can be completely covered by three non-overlapping squares is $3\sqrt{2}/4 \doteq 1.061$.

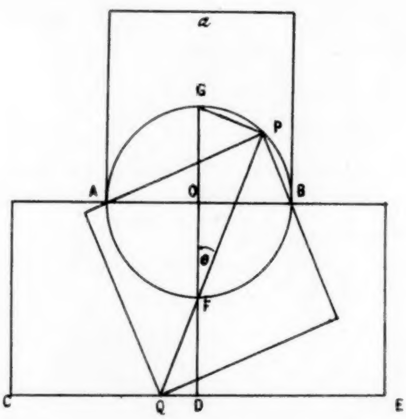


Figure 2

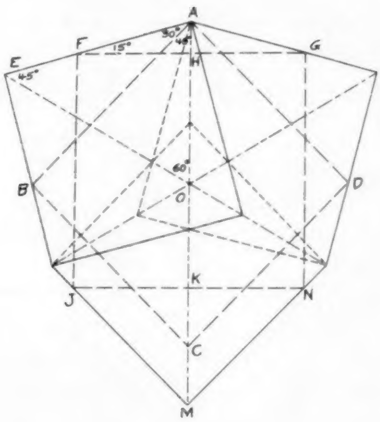


Figure 3

Wilderman also offered a symmetrical configuration in which vertices of the three overlapping squares coincided and the diagonals made angles of 120° with each other as in Figure 3. This arrangement contains two relatively maximum squares. The diagonal of one, $ABCD$, falls along a diagonal of one of the unit squares. It has a side $AB = AE \sec 30°$ $= 2\sqrt{3} \doteq 1.155$ as derived from triangle AEB. The computation of the side of the other square, $FGNJ$, which has sides parallel to one of the unit square's diagonals requires more labor. $HK = FJ = s$, $AH =$ $(s/2)\tan 15°$, and $KM = KH = s/2$. From triangle AOE, $AO = \sin 45°/\sin 60°$, $OM = OE = \sin 75°/\sin 60°$. Now $AH + HK + KM = AO + OM$, so $(s/2)(\tan 15° + 3)$ $= (\sin 45° + \sin 75°)/\sin 60°$. Hence $s = (3\sqrt{6} + 4\sqrt{2})/11 \doteq 1.182$.

Curve Intersections with Integer Coordinates

96. [March 1951] *Proposed by R. E. Horton, Los Angeles City College.*

Determine conditions on the integers p and a such that both coordinates of the points of intersection of $x^2 + y^2 = a^2$ and $y^2 = 2px$ will be integers.

Solution by Leon Bankoff, Los Angeles, Calif. It is well-known that $x = 2mn$, $y = m^2 - n^2$, $a = m^2 + n^2$ and $x = m^2 - n^2$, $y = 2mn$, $a = m^2 + n^2$ (with m, n relatively prime and of different parity) give primitive solutions of $x^2 + y^2 = a^2$. Then $p = y^2/2x = (m^2 - n^2)^2/4mn$ or

$4m^2n^2/2(m^2 - n^2)$. But since p is an integer, we have as general solutions:

$$x = 8km^2n^2,\ y = 4kmn(m^2 - n^2),\ a = 4kmn(m^2 + n^2),\ p = k(m^2 - n^2)^2;$$

and

$$x = k(m^2 - n^2)^2,\ y = 2kmn(m^2 - n^2),\ a = k(m^4 - n^4),\ p = 2km^2n^2,$$

where k is an integer. Obviously, for every solution (a,p) there will be three other solutions $(a, -p)$, $(-a, p)$, and $(-a, -p)$.

Also solved by *L. A. Ringenberg, Eastern Illinois State College; and the proposer.*

To Construct a Right Triangle, given r and R

99. [May 1951] *Proposed by T. E. Sydnor, Pasadena City College, Calif.*

a) Construct a right triangle given the inradius and the circumradius. b) What conditions must be placed on r and R in order that the triangle shall have integer sides?

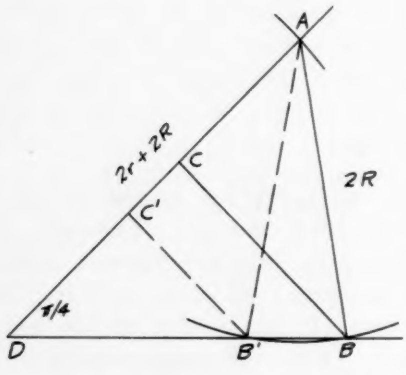


Figure 1

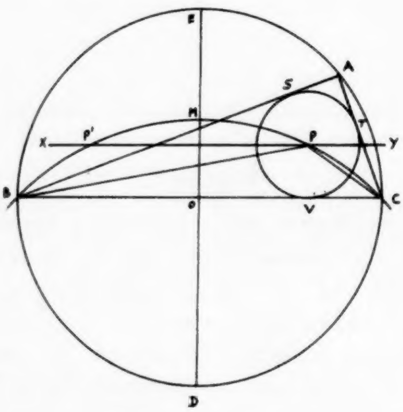


Figure 2

Solution by L. A. Ringenberg, Eastern Illinois State College. a) Construct an angle of magnitude $\pi/4$ and call its vertex D, as in Figure 1. On one side lay off $DA = 2R + 2r$. With center A and radius $2R$ describe a circle cutting the other side of the angle in B and B'. From B and B' drop perpendiculars BC and $B'C'$ to DA. Then BCA (or $B'C'A$) is the required triangle.

Obviously, circumradius of the triangle is R. It remains to show that its inradius is r. Let $AB = c$, $BC = a$, and $CA = b$. Since $DC' = C'B'$ and $DC = CB$, it follows that $a + b = 2r + 2R$, whence $a^2 + b^2 + 2ab = 4r(r + 2R) + 4R^2$. Now $a^2 + b^2 = c^2 = 4R^2$, so the area $K = ab/2 = r(r + 2R)$. But the area is also equal to the semiperimeter times the inradius. Therefore, since the semiperimeter is $r + 2R$, the inradius is r.

It is clear from the construction that there will be two, one, or

no solutions according as $R \gtreqless (1 + \sqrt{2})r$. The congruency of triangles BCA and $B'C'A$ is easily established.

We further observe that the sides a, b, c may be expressed in terms of r and R. $c = 2R$, and $a^2 = c^2 - b^2 = 4R^2 - (2r + 2R - a)^2$. It follows that a(or b) $= (R + r) \pm \sqrt{R^2 - 2Rr - r^2}$.

b) The last expression implies certain restrictions on r and R in order that the sides may be integers. However, it is more expeditious to employ the well-known parametric solution $a = 2wuv$, $b = w(u^2 - v^2)$, $c = w(u^2 + v^2)$, whereupon $R = c/2 = w(u^2 + v^2)/2$ and $r = (a + b - c)/2 = wv(u - v)$. Here u, v, w are integers with $u > v$ and u and v relatively prime and not both even.

Also solved by *Leon Bankoff, Los Angeles, California and the proposer,* both of whom used the construction method of Figure 2. There, two perpendicular diameters BC and ED were drawn in a circle of radius R. A second circle was described with center D and radius DB. Parallel to BC and at a distance r from DB a line XY cuts the second circle in P and P' the incenters. Tangents from B and C to the incircle complete the construction.

An Illegitimate Operation

103. [May 1951] *Proposed by M. S. Klamkin, Polytechnic Institute Brooklyn, N.Y.*

A high-school student solved the linear differential equation $dy/dx + Py = Q$ for y as if it were an ordinary algebraic equation. Under what conditions could this procedure have yielded a correct solution of the differential equation?

I. *Solution by P. B. Jordain, Columbia University, New York.* Solving

$$dy/dx + Py = Q \tag{1}$$

as an ordinary algebraic equation, the high-school student obviously cancelled the d's and got

$$y = Qx/(1 + Px) \tag{2}$$

with the condition that $P \neq -1/x$.

If (2) is to be a solution of (1), it must satisfy (1) identically. From (2) we obtain

$$y' = (Q + Q'x + PQ'x^2 - P'Qx^2)/(1 + Px)^2. \tag{3}$$

Substituting the values (2) and (3) in (1) and simplifying we get

$$Q'/Q = (P + P'x)/(1 + Px).$$

Thus

$$\log Q = \log k(1 + Px).$$

It follows that $y = Qx/(1 + Px)$ will be a solution of (1) if P and Q satisfy the general relation $Q = k(1 + Px)$.

II. *Solution by R. E. Winger, Los Angeles City College.* The high-school student presumably cancelled the d's, getting $y/x + Px = Q$. If this is to be a correct solution of the differential equation, then $y/x \equiv dy/dx$. That is, $y = cx$, an almost trivial result. When this is substituted into the original equation (in either form) we get $Q = c(1 + Px)$, the necessary relation between P and Q.

Also solved by *L. A. Ringenberg, Eastern Illinois State College; and the proposer.*

QUICKIES

From time to time this department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solution and the source, if known.

Q 48. Mentally multiply 38 by 32.

Q 49. n great circles on a sphere will, in general, intersect in $n(n - 1)$ points. Show how to place the numbers $1, 2, \cdots, n(n - 1)$ on these points in such a way that the sum of the numbers on every great circle is the same. [*Submitted by Leo Moser.*]

Q 50. Show that it is impossible to find integer solutions of $x^4 + y^4 + z^4 - x^4y^2 - y^4z^2 - z^4x^2 \pm x^2y^2z^2 = 0$. [*Emanuel Lasker in Scripta Mathematica,* **8**, 14, (Marxh 1941).]

ANSWERS

A 48. $38 \times 32 = (35 + 3)(35 - 3) = 35^2 - 3^2 = 1225 - 9 = 1216$.

A 49. Place any number a on any point and place its "complement", $b = n(n - 1) + 1 - a$, on the diametrically opposite point. Continue in this way until all the numbers (and all the points) are exhausted. Since each circle contains $n - 1$ pairs of diametrically opposite points, the numbers on each circle will total to $[n(n - 1) + 1](n - 1)$.

A 50. It may be assumed that x, y, z have no common factor hence are not all of the form $3k$. But if any or all of x, y, z are of the forms $3k \pm 1$, the left member is congruent to $\pm 1 \pmod 3$.

THE PERSONAL SIDE OF MATHEMATICS

This department desires especially articles showing what mathematics means to people in various professions and historical articles showing what classic mathematics meant to those who developed it. Material intended for this Department should be sent to the Mathematics Magazine, 14068 Van Nuys Blvd., Pacoima, California.

WHAT MATHEMATICS MEANS TO ME

Lewis Bayard Robinson

The author when a collegian at Johns Hopkins was haunted by that question:

What is truth?

And he knew that the motto of his college was:

"Veritas vos liberabit."

Therefore he resolved to become a mathematician because of all the sciences mathematics is the most rigorous and can best tell us what is truth.

The ancient Egyptians developed arithmetic and geometry for practical purposes. But this deeply religious people, as far as we know, did not see the mystical aspect of mathematics. The Babylonians, going a step further, were the best astronomers of antiquity coming very close to discovering the law of gravitation.[1] But from the time at least of Pythagoras, if not earlier, the Greeks saw the intimate connection of mathematics and metaphysics. And ever since the sixth century B.C., a century of great and rapid change, the metaphysical aspect of mathematics has never been completely lost to sight.

But all men are not mystics. The military engineer, important in the age of the two horned Alexander is, alas, important today. The practical sides of mathematics, applications to physics and other utilitarian sciences have tended to crowd out the metaphysical aspects.

The author is studying that profound work of Bergson:

"Durée et Simultanéité à propos de la Théorie d'Einstein".

It makes him feel that in the future mathematics will more and more enter the domain of metaphysics until it remakes that science as it has remade physics. Time, said Bergson, is non-homogeneous.[2] If unable to go to sleep at night the author likes to meditate on the possibility of constructing a mathematical theory where time and space are non-homogeneous. Now space-time is a multitude of simultaneous events, in number infinite, and each flowing with a different velocity. Perhaps the study of functions depending on variables, infinite in number, will set us on the road to discovery.

[1] Tarn—The Greeks in Bactria and India, p. 44.

[2] Time and Free Will.

Baltimore, Maryland